

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 08:15:07 ; Search time 36 Seconds
(without alignments)
70.327 Million cell updates/sec

Title: US-09-687-993-18

Perfect score: 104

Sequence: 1 NPENSRGKGRGQKMRG 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

ched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

- 1: /SID82/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SID82/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SID82/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SID82/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SID82/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SID82/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SID82/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SID82/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SID82/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SID82/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SID82/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SID82/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SID82/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SID82/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SID82/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SID82/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SID82/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SID82/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SID82/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SID82/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SID82/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SID82/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SID82/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	104	100.0	19	AAW15722	Truncated GDNF N-t
2	104	100.0	20	AAW15723	Truncated GDNF N-t
3	104	100.0	21	AAW15724	Truncated GDNF N-t
4	104	100.0	22	AAW15725	Truncated GDNF N-t
5	104	100.0	23	AAW15726	Truncated GDNF N-t
6	104	100.0	24	AAW15727	Truncated GDNF N-t
7	104	100.0	25	AAW15728	Truncated GDNF N-t
8	104	100.0	26	AAW15729	Truncated GDNF N-t
9	104	100.0	27	AAW15730	Truncated GDNF N-t
10	104	100.0	28	AAW15731	Truncated GDNF N-t

11	104	100.0	29	18	AAW15732	Truncated GDNF N-t
12	104	100.0	30	18	AAW15733	Truncated GDNF N-t
13	104	100.0	31	18	AAW15734	Truncated GDNF N-t
14	104	100.0	32	18	AAW15735	Truncated GDNF N-t
15	104	100.0	33	18	AAW15736	Truncated GDNF N-t
16	104	100.0	34	18	AAW15737	Truncated GDNF N-t
17	104	100.0	35	18	AAW15738	Truncated GDNF N-t
18	104	100.0	36	18	AAW15739	Truncated GDNF N-t
19	104	100.0	37	18	AAW15740	Truncated GDNF N-t
20	104	100.0	38	18	AAW15741	Truncated GDNF N-t
21	104	100.0	39	18	AAW15742	Truncated GDNF N-t
22	104	100.0	133	16	AAW15736	Human ATF-2. Homo
23	104	100.0	133	18	AAW30069	Human glial-derive
24	104	100.0	134	18	AAW32106	Human partial glia
25	104	100.0	134	18	AAW18052	Mature human glial
26	104	100.0	134	18	AAW18058	Mature human glial
27	104	100.0	134	18	AAW22027	Glial cell derived
28	104	100.0	134	18	AAW23782	Human mature glial
29	104	100.0	134	18	AAW15706	Glial cell line-de
30	104	100.0	134	20	AAW16658	WO9914235 Seq ID N
31	104	100.0	134	22	AAW35941	GDNF amino acid se
32	104	100.0	134	23	AAW51947	Human TGPbeta prot
33	104	100.0	135	18	AAW14930	Human glial cell 1
34	104	100.0	135	18	AAW31945	Human glial cell 1
35	104	100.0	159	20	AAW50698	Human GDNF protein
36	104	100.0	160	14	AAW38298	Human glial derive
37	104	100.0	185	16	AAW79375	Human ATF-1. Homo
38	104	100.0	185	20	AAW50697	Human GDNF protein
39	104	100.0	211	19	AAW83964	Human glial cell 1
40	104	100.0	211	20	AAW50695	Human GDNF protein
41	104	100.0	211	20	AAW50696	Human GDNF protein
42	99	95.2	134	20	AAW16659	WO9914235 Seq ID N
43	99	95.2	134	20	AAW16660	WO9914235 Seq ID N
44	99	95.2	134	22	AAU03952	Rat GDNF polypepti
45	99	95.2	147	22	AAU04452	Rat mutant G-hf-GD

ALIGNMENTS

- RESULT 1
- AAW15722
- ID AAW15722 standard; Peptide; 19 AA.
- XX
- AC AAW15722;
- XX
- DF 28-NOV-1997 (first entry)
- XX
- DE Truncated GDNF N-terminal peptide.
- DE
- XX
- KW Glial cell line-derived neurotrophic factor; GDNF; human; dopaminergic; nerve cell; Parkinson's disease; gene therapy.
- KW
- XX
- OS Synthetic.
- XX
- XX
- PN WO9711964-A1.
- XX
- PD 03-APR-1997.
- XX
- PF 16-SEP-1996; 96WO-US14915.
- XX
- PR 28-SEP-1995; 95US-0535681.
- XX
- PA (AMGE-) AMGEN INC.
- XX
- PI Hu SS;
- XX
- DR WPI; 1997-212849/19.
- XX
- PT Truncated glial cell line-derived neurotrophic factor protein - used in the treatment and gene therapy of Parkinson's disease
- PT
- XX
- PS Claim 1; Page 83; 105pp; English.

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NPENSRGKRRGGRGKNGR 19
 {}
 Db 3 NPENSRGKRRGGRGKNGR 21

RESULT 4
 AAW15725
 ID AAW15725 standard; Peptide; 22 AA.
 AC AAW15725;
 XX
 XX 28-NOV-1997 (first entry)
 XX
 DE Truncated GDNF N-terminal peptide.
 XX
 KW Glial cell line-derived neurotrophic factor; GDNF; human;
 KW dopaminergic; nerve cell; Parkinson's disease; gene therapy.
 US Synthetic.
 XX
 XX WO9711964-A1.
 PN
 XX
 XX 03-APR-1997.
 PD
 XX 16-SEP-1996; 96WO-US14915.
 PF
 XX 28-SEP-1995; 95US-0535681.
 PR
 XX (AMGE-) AMGEN INC.
 PA
 XX Hu SS;
 PI
 XX WPI; 1997-212849/19.
 DR
 XX Truncated glial cell line-derived neurotrophic factor protein - used
 PT in the treatment and gene therapy of Parkinson's disease
 PT
 XX Claim 1; Page 84; 105pp; English.
 PS
 XX This peptide sequence comprises amino acid residues 19-40 of the
 CC human glial cell line-derived neurotrophic factor (GDNF) protein
 CC (see AAW15706) and represents an N-terminal sequence of a novel
 CC truncated GDNF. Claimed truncated GDNF proteins have the formula:
 CC X-(Cys41-Cys133)-Y, where (Cys41-Cys133) = Cys41 through Cys133 of
 CC mature human GDNF; Y = a C-terminal Cys133 or Ile134; and X = a
 CC methionylated or nonmethionylated amine group of Cys41 or an
 CC N-terminus selected from G, RG, NRG or the N-terminal peptides
 CC given in AAW15707-42, and additions, substitutions and internal
 CC deletion variants of these. Also claimed are: a polynucleotide
 CC (see AAT60542-46) encoding a truncated GDNF (see AAW15743-45); a vector;
 CC a transformed or transfected prokaryotic or eukaryotic host cell;
 CC and a GDNF composition comprising mature GDNF protein (44 kDa) and
 CC one or more truncated GDNFs (29-40 kDa). The truncated GDNF is used
 CC in the treatment of nervous system damage caused by disease or
 CC injury, especially in the treatment of Parkinson's disease.
 XX
 XX Sequence 22 AA;
 SQ

Query Match 100.0%; Score 104; DB 18; Length 22;
 Best Local Similarity 100.0%; Pred. No. 6.5e-08;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NPENSRGKRRGGRGKNGR 19
 {}
 Db 4 NPENSRGKRRGGRGKNGR 22

RESULT 5
 AAW15726
 ID AAW15726 standard; Peptide; 23 AA.
 XX

AC AAW15726;
 XX
 DT 28-NOV-1997 (first entry)
 XX
 DE Truncated GDNF N-terminal peptide.
 XX
 KW Glial cell line-derived neurotrophic factor; GDNF; human;
 KW dopaminergic; nerve cell; Parkinson's disease; gene therapy.
 XX
 OS Synthetic.
 XX
 XX WO9711964-A1.
 PN
 XX
 XX 03-APR-1997.
 PD
 XX 16-SEP-1996; 96WO-US14915.
 PF
 XX 28-SEP-1995; 95US-0535681.
 PR
 XX (AMGE-) AMGEN INC.
 PA
 XX Hu SS;
 PI
 XX WPI; 1997-212849/19.
 DR
 XX Truncated glial cell line-derived neurotrophic factor protein - used
 PT in the treatment and gene therapy of Parkinson's disease
 PT
 XX Claim 1; Page 84; 105pp; English.
 PS
 XX This peptide sequence comprises amino acid residues 18-40 of the
 CC human glial cell line-derived neurotrophic factor (GDNF) protein
 CC (see AAW15706) and represents an N-terminal sequence of a novel
 CC truncated GDNF. Claimed truncated GDNF proteins have the formula:
 CC X-(Cys41-Cys133)-Y, where (Cys41-Cys133) = Cys41 through Cys133 of
 CC mature human GDNF; Y = a C-terminal Cys133 or Ile134; and X = a
 CC methionylated or nonmethionylated amine group of Cys41 or an
 CC N-terminus selected from G, RG, NRG or the N-terminal peptides
 CC given in AAW15707-42, and additions, substitutions and internal
 CC deletion variants of these. Also claimed are: a polynucleotide
 CC (see AAT60542-46) encoding a truncated GDNF (see AAW15743-45); a vector;
 CC a transformed or transfected prokaryotic or eukaryotic host cell;
 CC and a GDNF composition comprising mature GDNF protein (44 kDa) and
 CC one or more truncated GDNFs (29-40 kDa). The truncated GDNF is used
 CC in the treatment of nervous system damage caused by disease or
 CC injury, especially in the treatment of Parkinson's disease.
 XX
 XX Sequence 23 AA;
 SQ

Query Match 100.0%; Score 104; DB 18; Length 23;
 Best Local Similarity 100.0%; Pred. No. 6.8e-08;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NPENSRGKRRGGRGKNGR 19
 {}
 Db 5 NPENSRGKRRGGRGKNGR 23

RESULT 6
 AAW15727
 ID AAW15727 standard; Peptide; 24 AA.
 XX
 AC AAW15727;
 XX
 XX 28-NOV-1997 (first entry)
 DT
 XX Truncated GDNF N-terminal peptide.
 DE
 XX Glial cell line-derived neurotrophic factor; GDNF; human;
 KW dopaminergic; nerve cell; Parkinson's disease; gene therapy.
 XX
 XX Synthetic.
 XX

```
PN WO9711964-A1.
XX
PD 03-APR-1997.
XX
PF 16-SEP-1996; 96WO-US14915.
XX
PR 28-SEP-1995; 95US-0535681.
XX
XX (AMGE-) AMGEN INC.
PA
XX Hu SS;
XX
PI
XX
DR WPI; 1997-212849/19.
XX
PT Truncated glial cell line-derived neurotrophic factor protein - used
PT in the treatment and gene therapy of Parkinson's disease
XX
PS Claim 1; Page 84; 105pp; English.
XX
CC This peptide sequence comprises amino acid residues 17-40 of the
CC human glial cell line-derived neurotrophic factor (GDNF) protein
CC (see AAW15706) and represents an N-terminal sequence of a novel
CC truncated GDNF. Claimed truncated GDNF proteins have the formula:
CC X-(Cys41-Cys133)-Y, where (Cys41-Cys133) = Cys41 through Cys133 of
CC mature human GDNF; Y = a C-terminal Cys133 or Ile134; and X = a
CC methionylated or nonmethionylated amine group of Cys41 or an
CC N-terminus selected from G, RG, NRG or the N-terminal peptides
CC given in AAW15707-42, and additions, substitutions and internal
CC deletion variants of these. Also claimed are: a polynucleotide
CC (see AAT60542-46) encoding a truncated GDNF (see AAW15743-45); a vector;
CC a transformed or transfected prokaryotic or eukaryotic host cell;
CC and a GDNF composition comprising mature GDNF protein (44 kDa) and
CC one or more truncated GDNFs (29-40 kDa). The truncated GDNF is used
CC in the treatment of nervous system damage caused by disease or
CC injury, especially in the treatment of Parkinson's disease.
XX
SQ Sequence 24 AA;
XX
Query Match 100.0%; Score 104; DB 18; Length 24;
Best Local Similarity 100.0%; Pred. No. 7.1e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NPENSRGKRRGQGRGNRG 19
Db |||||
6 NPENSRGKRRGQGRGNRG 24
RESULT 7
AAW15728
AAW15728 standard; Peptide; 25 AA.
AAW15728;
XX
XX 28-NOV-1997 (first entry)
XX
DE Truncated GDNF N-terminal peptide.
XX
KW Glial cell line-derived neurotrophic factor; GDNF; human;
KW dopaminergic; nerve cell; Parkinson's disease; gene therapy.
XX
OS Synthetic.
XX
XX WO9711964-A1.
PN
XX 28-NOV-1997 (first entry)
XX
DE Truncated GDNF N-terminal peptide.
XX
KW Glial cell line-derived neurotrophic factor; GDNF; human;
KW dopaminergic; nerve cell; Parkinson's disease; gene therapy.
XX
OS Synthetic.
XX
XX WO9711964-A1.
PN
XX 03-APR-1997.
XX
XX 16-SEP-1996; 96WO-US14915.
XX
XX 28-SEP-1995; 95US-0535681.
XX
XX (AMGE-) AMGEN INC.
PA
XX Hu SS;
XX
PI
XX
DR WPI; 1997-212849/19.
XX
PT Truncated glial cell line-derived neurotrophic factor protein - used
PT in the treatment and gene therapy of Parkinson's disease
XX
PS Claim 1; Page 84; 105pp; English.
XX
CC This peptide sequence comprises amino acid residues 15-40 of the
CC human glial cell line-derived neurotrophic factor (GDNF) protein
CC (see AAW15706) and represents an N-terminal sequence of a novel
CC truncated GDNF. Claimed truncated GDNF proteins have the formula:
CC X-(Cys41-Cys133)-Y, where (Cys41-Cys133) = Cys41 through Cys133 of
CC mature human GDNF; Y = a C-terminal Cys133 or Ile134; and X = a
CC methionylated or nonmethionylated amine group of Cys41 or an
CC N-terminus selected from G, RG, NRG or the N-terminal peptides
CC given in AAW15707-42, and additions, substitutions and internal
CC deletion variants of these. Also claimed are: a polynucleotide
CC (see AAT60542-46) encoding a truncated GDNF (see AAW15743-45); a vector;
CC a transformed or transfected prokaryotic or eukaryotic host cell;
CC and a GDNF composition comprising mature GDNF protein (44 kDa) and
CC one or more truncated GDNFs (29-40 kDa). The truncated GDNF is used
CC in the treatment of nervous system damage caused by disease or
CC injury, especially in the treatment of Parkinson's disease.
XX
SQ Sequence 25 AA;
XX
Query Match 100.0%; Score 104; DB 18; Length 25;
Best Local Similarity 100.0%; Pred. No. 7.3e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NPENSRGKRRGQGRGNRG 19
Db |||||
7 NPENSRGKRRGQGRGNRG 25
RESULT 8
AAW15729
AAW15729 standard; Peptide; 26 AA.
XX
XX 28-NOV-1997 (first entry)
XX
DE Truncated GDNF N-terminal peptide.
XX
KW Glial cell line-derived neurotrophic factor; GDNF; human;
KW dopaminergic; nerve cell; Parkinson's disease; gene therapy.
XX
OS Synthetic.
XX
XX WO9711964-A1.
PN
XX 03-APR-1997.
XX
XX 16-SEP-1996; 96WO-US14915.
XX
XX 28-SEP-1995; 95US-0535681.
XX
XX (AMGE-) AMGEN INC.
PA
XX Hu SS;
XX
PI
XX
DR WPI; 1997-212849/19.
XX
PT Truncated glial cell line-derived neurotrophic factor protein - used
PT in the treatment and gene therapy of Parkinson's disease
XX
PS Claim 1; Page 84; 105pp; English.
XX
CC This peptide sequence comprises amino acid residues 15-40 of the
CC human glial cell line-derived neurotrophic factor (GDNF) protein
CC (see AAW15706) and represents an N-terminal sequence of a novel
```

CC truncated GDNF. Claimed truncated GDNF proteins have the formula:
CC X-(Cys41-Cys133)-Y, where (Cys41-Cys133) = Cys41 through Cys133 of
CC mature human GDNF; Y = a C-terminal Cys133 or Ile134; and X = a
CC methionylated or nonmethionylated amine group of Cys41 or an
CC N-terminus selected from G, RG, NRG or the N-terminal peptides
CC given in AAW15707-42, and additions, substitutions and internal
CC deletion variants of these. Also claimed are: a polynucleotide
CC (see AAT60542-46) encoding a truncated GDNF (see AAW15743-45); a vector;
CC a transformed or transfected prokaryotic or eukaryotic host cell;
CC and a GDNF composition comprising mature GDNF protein (44 kDa) and
CC one or more truncated GDNFs (29-40 kDa). The truncated GDNF is used
CC in the treatment of nervous system damage caused by disease or
CC injury, especially in the treatment of Parkinson's disease.

XX SQ Sequence 26 AA;

Query Match 100.0%; Score 104; DB 18; Length 26;
Best Local Similarity 100.0%; Pred. No. 7.6e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NPENSRGKGRGGRGKNGR 19
|||||
DB 8 NPENSRGKGRGGRGKNGR 26

RESULT 9
AAW15730
ID AAW15730 standard; Peptide; 27 AA.

XX AC AAW15730;

XX DT 28-NOV-1997 (first entry)

XX DE Truncated GDNF N-terminal peptide.

XX KW Glial cell line-derived neurotrophic factor; GDNF; human;
XX KQ dopaminergic; nerve cell; Parkinson's disease; gene therapy.

XX OS Synthetic.

XX PN WO9711964-A1.

XX PD 03-APR-1997.

XX PF 16-SEP-1996; 96WO-US14915.

XX PS 28-SEP-1995; 95US-0535681.

XX (AMGE-) AMGEN INC.

XX PI Hu SS;

XX DR WPI; 1997-212849/19.

XX PT Truncated glial cell line-derived neurotrophic factor protein - used
XX in the treatment and gene therapy of Parkinson's disease

XX PS Claim 1; Page 84; 105pp; English.

XX CC This peptide sequence comprises amino acid residues 14-40 of the
CC human glial cell line-derived neurotrophic factor (GDNF) protein
CC (see AAW15706) and represents an N-terminal sequence of a novel
CC truncated GDNF. Claimed truncated GDNF proteins have the formula:
CC X-(Cys41-Cys133)-Y, where (Cys41-Cys133) = Cys41 through Cys133 of
CC mature human GDNF; Y = a C-terminal Cys133 or Ile134; and X = a
CC methionylated or nonmethionylated amine group of Cys41 or an
CC N-terminus selected from G, RG, NRG or the N-terminal peptides
CC given in AAW15707-42, and additions, substitutions and internal
CC deletion variants of these. Also claimed are: a polynucleotide
CC (see AAT60542-46) encoding a truncated GDNF (see AAW15743-45); a vector;
CC a transformed or transfected prokaryotic or eukaryotic host cell;
CC and a GDNF composition comprising mature GDNF protein (44 kDa) and
CC one or more truncated GDNFs (29-40 kDa). The truncated GDNF is used

CC in the treatment of nervous system damage caused by disease or
CC injury, especially in the treatment of Parkinson's disease.

XX SQ Sequence 27 AA;

Query Match 100.0%; Score 104; DB 18; Length 27;
Best Local Similarity 100.0%; Pred. No. 7.9e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NPENSRGKGRGGRGKNGR 19
|||||
DB 9 NPENSRGKGRGGRGKNGR 27

RESULT 10

AAW15731

ID AAW15731 standard; Peptide; 28 AA.

XX AC AAW15731;

XX DT 28-NOV-1997 (first entry)

XX DE Truncated GDNF N-terminal peptide.

XX KW Glial cell line-derived neurotrophic factor; GDNF; human;
XX KQ dopaminergic; nerve cell; Parkinson's disease; gene therapy.

XX OS Synthetic.

XX PN WO9711964-A1.

XX PD 03-APR-1997.

XX PF 16-SEP-1996; 96WO-US14915.

XX PR 28-SEP-1995; 95US-0535681.

XX PA (AMGE-) AMGEN INC.

XX PI Hu SS;

XX DR WPI; 1997-212849/19.

XX PT Truncated glial cell line-derived neurotrophic factor protein - used
XX in the treatment and gene therapy of Parkinson's disease

XX PS Claim 1; Page 84; 105pp; English.

XX CC This peptide sequence comprises amino acid residues 13-40 of the
CC human glial cell line-derived neurotrophic factor (GDNF) protein
CC (see AAW15706) and represents an N-terminal sequence of a novel
CC truncated GDNF. Claimed truncated GDNF proteins have the formula:
CC X-(Cys41-Cys133)-Y, where (Cys41-Cys133) = Cys41 through Cys133 of
CC mature human GDNF; Y = a C-terminal Cys133 or Ile134; and X = a
CC methionylated or nonmethionylated amine group of Cys41 or an
CC N-terminus selected from G, RG, NRG or the N-terminal peptides
CC given in AAW15707-42, and additions, substitutions and internal
CC deletion variants of these. Also claimed are: a polynucleotide
CC (see AAT60542-46) encoding a truncated GDNF (see AAW15743-45); a vector;
CC a transformed or transfected prokaryotic or eukaryotic host cell;
CC and a GDNF composition comprising mature GDNF protein (44 kDa) and
CC one or more truncated GDNFs (29-40 kDa). The truncated GDNF is used
CC in the treatment of nervous system damage caused by disease or
CC injury, especially in the treatment of Parkinson's disease.

XX SQ Sequence 28 AA;

Query Match 100.0%; Score 104; DB 18; Length 28;
Best Local Similarity 100.0%; Pred. No. 8.1e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NPENSRGKGRGGRGKNGR 19
|||||

in the treatment and gene therapy of Parkinson's disease

Claim 1: Page 84: 105pp: English.

This peptide sequence comprises amino acid residues 9-40 of the human glial cell line-derived neurotrophic factor (GDNF) protein (see AA015706) and represents an N-terminal sequence of a novel truncated GDNF. Claimed truncated GDNF proteins have the formula: X-(Cys41-Cys133)-Y, where (Cys41-Cys133) = Cys41 through Cys133 of mature human GDNF; Y = a C-terminal Cys133 or Ile134; and X = a methionylated or nonmethionylated amine group of Cys41 or an N-terminus selected from G, RG, NRG or the N-terminal peptides given in AA015707-42, and additions, substitutions and internal deletion variants of these. Also claimed are: a polynucleotide (see AA060542-46) encoding a truncated GDNF (see AA015743-45); a transformed or transfected prokaryotic or eukaryotic host cell; and a GDNF composition comprising mature GDNF protein (44 kDa) and one or more truncated GDNFs (29-40 kDa). The truncated GDNF is used in the treatment of nervous system damage caused by disease or injury, especially in the treatment of Parkinson's disease.

Sequence 32 AA;

Query Match 100.0%; Score 104; DB 18; Length 32;
Best Local Similarity 100.0%; Pred. No. 9,2e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 15
AAW15736
ID AAW15736 standard; Peptide: 33 AA.
XX
XX AAW15736;
XX
XX
XX
XX 28-NOV-1997 (first entry)
XX
XX Truncated GDNF N-terminal peptide.
DE

KW	glial cell line-derived neurotrophic factor; GDNF; human;
KK	dopaminergic; nerve cell; Parkinson's disease; gene therapy.
OS	Synthetic.
XX	
XX	
PN	W09711964-A1.
XX	
PD	03-APR-1997.
XX	
PD	
PF	16-SEP-1996; 96WO-US14915.
XX	
XX	
PR	28-SEP-1995; 95US-0535681.
XX	
PA	{AMGE-} AMGEN INC.
XX	
PI	Hu SS;
XX	
DR	WPI; 1997-212849/19.
XX	
PT	Truncated glial cell line-derived neurotrophic factor protein - used
PT	in the treatment and gene therapy of Parkinson's disease
XX	
PS	Claim 1; Page 84; 105pp; English.
XX	
CC	This peptide sequence comprises amino acid residues 8-40 of the
CC	human glial cell line-derived neurotrophic factor (GDNF) protein
CC	(see AAU15706) and represents an N-terminal sequence of a novel
CC	truncated GDNF. Claimed truncated GDNF proteins have the formula:
CC	X-(Cys41-Cys133)-Y, where (Cys41-Cys133) = Cys41 through Cys133 of
CC	mature human GDNF; Y = a C-terminal Cys133 or Ile134; and X = a
CC	methylated or nonmethylated amine group of Cys41 or an

CC N-terminus selected from G, RG, NRG or the N-terminal peptides
CC given in AAW15707-42, and additions, substitutions and internal
CC deletion variants of these. Also claimed are: a polynucleotide
CC (see AAT60542-46) encoding a truncated GDNF (see AAW15743-45); a vector;
CC a transfected or transfected prokaryotic or eukaryotic host cell;
CC and a GDNF composition comprising mature GDNF protein (44 kDa) and
CC one or more truncated GDNFs (29-40 kDa). The truncated GDNF is used
CC in the treatment of nervous system damage caused by disease or
CC injury, especially in the treatment of Parkinson's disease.
XX
SQ Sequence 33 AA;

Query Match 100.0%; Score 104; DB 18; Length 33;
Best Local Similarity 100.0%; Pred. No. 9.5e-08;
Matches 19: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NPENSRGKRRGQGGKNRG 19
DB 15 NPENSRGKRRGQGGKNRG 33

rch completed: December 4, 2002, 08:17:58
time : 37 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 08:17:02 ; Search time 16 Seconds
(without alignments)
114.160 Million cell updates/sec

Title: US-09-687-993-18

Perfect score: 104

Sequence: 1 NPENSRGKRRGQGNKNG 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

al number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	104	100.0	211	2 B37499	glial cell line-de
2	99	95.2	211	2 A37499	glial cell line-de
3	99	95.2	211	2 I49686	glial cell line-de
4	55.5	53.4	2925	2 T00133	RNA-directed RNA p
5	54	51.9	898	2 H87481	ribonuclease, Rne/
6	51	49.0	525	2 B48058	RNA-binding protei
7	51	49.0	750	2 A97501	topoisomerase iv c
8	51	49.0	750	2 AE2719	topoisomerase IV s
9	51	49.0	1334	2 E86451	probable copia-ty
10	50	48.1	340	1 WMBEL1	latency-related pr
11	50	48.1	613	2 F64056	probable ATP-depen
12	49.5	47.6	296	2 T01458	hypothetical prote
13	49.5	47.6	893	2 AH1867	hypothetical prote
14	49	47.1	166	2 T39586	rna binding protei
15	49	47.1	494	2 S36541	E2 protein - human
16	48.5	46.6	637	2 G88533	70.5K hypothetical
17	48.5	46.6	633	2 T02673	heterogeneous nucl
18	48	46.2	126	2 T16952	hypothetical prote
19	47.5	45.7	117	2 T38440	small nuclear ribo
20	47.5	45.7	523	2 E96576	unknown protein, 4
21	47.5	45.7	695	2 I51552	dsRNA-binding prot
22	47.5	45.7	1339	2 T47841	hypothetical prote
23	47	45.2	269	2 JCT7700	38K ribosome-assoc
24	47	45.2	300	2 T49225	hypothetical prote
25	47	45.2	536	2 T42606	probable transcrip
26	47	45.2	656	2 D96831	hypothetical prote
27	47	45.2	849	2 A96592	hypothetical prote
28	47	45.2	1171	2 T05039	hypothetical prote
29	46.5	44.7	800	2 I51653	dsRNA-binding prot

ALIGNMENTS

RESULT 1

B37499
glial cell line-derived neurotrophic factor precursor - human
N:Alternate names: GDNF
C:Species: Homo sapiens (man)
C:Date: 26-Aug-1999 #sequence_revision 26-Aug-1999 #text_change 26-Aug-1999
C:Accession: B37499
R:Lin, L.F.; Doherty, D.H.; Lile, J.D.; Bektesh, S.; Collins, F.
Science 260, 1130-1132, 1993
A:Title: GDNF: a glial cell line-derived neurotrophic factor for midbrain dopaminergi
A:Reference number: A37499; MUID:93262463; PMID:8493557
A:Accession: B37499
A:Molecule type: DNA
A:Residues: 1-211 <LIN>
A:Cross-references: GB:I15306; GB:I15306; MID:g306761; PIDN:AAA67910.1; PID:g306763
A:Note: sequence extracted from NCBI backbone (NCBIP:132084)
C:Keywords: glycoprotein; homodimer
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-77/Domain: propeptide #status predicted <PRO>
F:78-211/Product: glial cell line-derived neurotrophic factor #status predicted <MAT>
F:126,162/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 104; DB 2; Length 211;

Best Local Similarity 100.0%; Pred. No. 1.7e-07;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NPENSRGKRRGQGNKNG 19

|||||

Db 99 NPENSRGKRRGQGNKNG 117

RESULT 2

A37499
glial cell line-derived neurotrophic factor precursor - rat
N:Alternate names: GDNF
N:Contains: Rattus norvegicus (Norway rat)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 16-Feb-1994 #sequence_revision 16-Feb-1994 #text_change 05-Nov-1999
C:Accession: A37499; I57605; I53427; I58180; S61537
R:Lin, L.F.; Doherty, D.H.; Lile, J.D.; Bektesh, S.; Collins, F.
Science 260, 1130-1132, 1993
A:Title: GDNF: a glial cell line-derived neurotrophic factor for midbrain dopaminergi
A:Reference number: A37499; MUID:93262463; PMID:8493557
A:Accession: A37499
A:Molecule type: mRNA; protein
A:Residues: 1-211 <LIN>
A:Cross-references: GB:I15305; MID:g310123; PIDN:AAA67909.1; PID:g310124
A:Experimental source: glial cell line B49
A:Note: sequence extracted from NCBI backbone (NCBIP:132083)
R:Springer, J.E.; Seeburger, J.L.; He, J.; Gabrea, A.; Blankenhorn, E.P.; Bergman, L.
Exp. Neurol. 131, 47-52, 1995
A:Title: cDNA sequence and differential mRNA regulation of two forms of glial cell li

RESULT 4

T00133
RNA-directed RNA polymerase (EC 2.7.7.48) - rice grassy stunt virus
C:Species: rice grassy stunt virus
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-Jul-2000
C:Accession: T00133
R:Toriyama, S.; Kimishima, T.; Takahashi, M.; Shimizu, T.; Minaka, N.; Akutsu, K.
J. Gen. Virol. 79, 2051-2058, 1998
A:Title: The complete nucleotide sequence of the rice grassy stunt virus genome and g
A:Reference number: 214118; MUID: 98378059; PMID: 9714257
A:Accession: T00133
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: genomic RNA
A:Residues: 1-2925 <TOR>
A:Cross-references: EMBL:AB009656; NID:g3410897; PIDN:BAA32246.1; PID:g3410899
A:Experimental source: isolate IRRI
C:Keywords: nucleotidyltransferase

Query Match 53.4%; Score 55.5; DB 2; Length 2925;
Best Local Similarity 64.7%; Pred. No. 11;
Matches 11; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

Qy 1 NPENSRGKGRGGRGKN 17
||:||||:|||||:|::

Db 2907 NPPSRGRGRRG-RGRS 2922

RESULT 5

H87481
Ribonuclease, Rne/Rng family protein [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: H87481
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Neilson, K.E.; Eisen, J.; Heidelberg,
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: AB7249; MUID: 21173698; PMID: 11259647
A:Accession: H87481
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-898 <STO>
A:Cross-references: GB:AE005673; NID:g13423322; PIDN:AAK23852.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC1877

Query Match 51.9%; Score 54; DB 2; Length 898;
Best Local Similarity 57.9%; Pred. No. 6.5;
Matches 11; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

Qy 3 ENSRG--KGRGQRGKNRG 19
::|||:|||||:|::|

Db 604 DDERGDRKGRGRDRNRG 622

RESULT 6

B48058
RNA-binding protein NAB2 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein G2910; protein YGL122c
C:Species: Saccharomyces cerevisiae
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 21-Jul-2000
C:Accession: B48058; A48058; S64132
R:Anderson, J.T.; Wilson, S.M.; Datar, K.V.; Swanson, M.S.
Mol. Cell. Biol. 13, 2730-2741, 1993
A:Title: NAB2: a yeast nuclear polyadenylated RNA-binding protein essential for cell
A:Reference number: A48058; MUID: 93233636; PMID: 8474438
A:Accession: B48058
A:Molecule type: DNA
A:Residues: 1-525 <AND>
A:Cross-references: CB:110298; NID:C295528; PIDN:AAJ4819.1; RID:C295528

A:Experimental source: strain BJ926
 A:Accession: A48058
 A:Molecule type: DNA
 A:Residues: 1-120,149-525 <ANN>
 A:Cross-references: NID:g295630; PID:AAA34820.1; PID:g295631
 A:Experimental source: strain YNN318
 A:Note: sequence extracted from NCBI backbone (NCBIN:129803, NCBIP:129804)
 R:Laquin, G.
 submitted to the Protein Sequence Database, May 1996
 A:Reference number: S64122
 A:Accession: S64132
 A:Molecule type: DNA
 A:Residues: 1-525 <LAU>
 A:Cross-references: EMBL:272644; NID:g1322680; PID:e243350; PID:g1322681; MIPS:YGL122c
 A:Experimental source: strain S288C
 C:Genetics:
 A:Gene: SGD:NAB2
 A:Cross-references: SGD:S0003090; MIPS:YGL122c
 A:Map position: 7L
 Keywords: nucleus; RNA binding

Query Match 49.0%; Score 51; DB 2; Length 525;
 Best Local Similarity 76.9%; Pred. No. 11;
 Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 7 GKRRGGRGQRKNG 19
 || ||| || |||
 DB 215 GKNRRGGRGNG 227

RESULT 7
 A97501
 topoisomerase iv chain a [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
 C:Species: Agrobacterium tumefaciens
 C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
 Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
 A:Reference number: A97359; PMID:11743194
 A:Accession: A97501
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-750 <KUR>
 A:Cross-references: GB:AE007869; PID:AAK8562.1; PID:g15156198; GSPDB:GN00169
 C:Genetics:
 A:Gene: AGR_C_2144
 A:Map position: circular chromosome
 C:Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisomerase

Query Match 49.0%; Score 51; DB 2; Length 750;
 Best Local Similarity 66.7%; Pred. No. 15;
 Matches 12; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 PENSRRGGRGQRKNG 19
 || ||| || |||
 DB 677 PENSRRGGRGQRKNG 694

RESULT 8
 AE2719
 topoisomerase IV subunit A parC [imported] - Agrobacterium tumefaciens (strain C58, Dupo
 C:Species: Agrobacterium tumefaciens
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002
 C:Accession: AE2719
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, H
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuttyavin, T.; Levy, R.; Li, M.; McClell
 ; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, E.W.
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:11743193
 A:Accession: AE2719
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-750 <KUR>
 A:Cross-references: GB:AE008688; PID:AA42171.1; PID:g1779560; GSPDB:GN00186
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: parC
 A:Map position: circular chromosome
 C:Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisomeras

Query Match 49.0%; Score 51; DB 2; Length 750;
 Best Local Similarity 66.7%; Pred. No. 15;
 Matches 12; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 PENSRRGGRGQRKNG 19
 || ||| || |||
 DB 677 PENSRRGGRGQRKNG 694

RESULT 9
 E86451
 probable copia-type polyprotein, 28768-32772 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: E86451
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia
 Rizzo, M.; Rowley, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: E86451
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1334 <STO>
 A:Cross-references: GB:AE005172; NID:g10092575; PIDN:AA612968.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1
 C:Superfamily: retrovirus-related polyprotein

Query Match 49.0%; Score 51; DB 2; Length 1334;
 Best Local Similarity 60.0%; Pred. No. 24;
 Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 5 SRGRRGGRGQRKNG 19
 ||| ||| ||| |||
 DB 228 SRGRRGGRGQRKNG 242

RESULT 10
 WMBEL1
 latency-related protein 1 - human herpesvirus 1 (strain F)
 C:Species: human herpesvirus 1
 A:Note: host Homo sapiens (man)
 C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 10-Sep-1999
 C:Accession: A33337
 R:Wechsler, S.L.; Nesburn, A.B.; Zwaagstra, J.; Ghiasi, H.
 Virology 168, 168-172, 1989
 A:Title: Sequence of the latency-related gene of herpes simplex virus type 1.
 A:Reference number: A94388; MUID:89085598; PMID:2535901
 A:Accession: A33337
 A:Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-340 <WEC>
 A:Cross-references: GB:J04323; NID:g330133; PIDN:AAA45799.1; PID:g330134
 C:Genetics:

A:Introns: 249/2
 C:Superfamily: herpesvirus latency-related protein 1
 C:Keywords: tandem repeat
 F:26-41/Region: 16-residue tandem repeat
 F:42-57/Region: 16-residue tandem repeat
 F:58-73/Region: 16-residue tandem repeat

Query Match 48.1%; Score 50; DB 1; Length 340;
 Best Local Similarity 50.0%; Pred. No. 10;
 Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 PENSCKGRRGQGRGNRG 19
 | ||| ||| ||| |||
 Db 305 PRGSRGRRGGRGGRGG 322

RESULT 11
 F64056
 Probable ATP-dependent RNA helicase - Haemophilus influenzae (strain Rd KW20)
 C:Species: Haemophilus influenzae
 C:Note: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 02-Feb-2001
 C:Accession: F64056
 R:Flieischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M. Science 269, 496-512, 1995
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.
 A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
 A:Reference number: A64000; MUID:95350630; PMID:7542800
 A:Accession: F64056
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-613 <TIGR>
 A:Cross-references: GB:U32709; GB:L42023; NID:g1573190; PIDN:AAC21900.1; PID:g1573195; T:151-156/Region: nucleotide-binding motif A (P-loop)
 C:Keywords: ATP; nucleotide binding; P-loop
 F:49-56/Region: nucleotide-binding motif A (P-loop)
 F:151-156/Region: nucleotide-binding motif B
 F:155-158/Region: DEAD motif

Query Match 48.1%; Score 50; DB 2; Length 613;
 Best Local Similarity 44.8%; Pred. No. 17;
 Matches 13; Conservative 3; Mismatches 1; Indels 12; Gaps 2;

QY 3 ENSRG-----KGR-----RGQRGNRG 19
 |||||
 Db 556 DNSRGSDDFNGRKRGGRGGRGRG 584

RESULT 12
 F158
 Hypothetical protein F2401.20 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 23-Mar-2001
 C:Accession: T01458
 R:Shinn, P.; Buehler, E.; Dewar, K.; Feng, J.; Kim, C.; Li, Y.; Sun, H.; Conway, A.; Coriell, A.; Ecker, J.R.
 A:Submitted to the EMBL Data Library, January 1998
 A:Description: Genomic sequence for Arabidopsis thaliana BAC F2401.
 A:Reference number: Z14211
 A:Accession: T01458
 A:Status: translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-296 <SH1>
 A:Cross-references: EMBL:AC003113; NID:g26899438; PID:g2781364; GSPDB:GN00059; ATSP:F2401
 C:Genetics:
 A:Gene: ATSP:F2401.20
 A:Map position: 1

Query Match 47.6%; Score 49.5; DB 2; Length 296;
 Best Local Similarity 68.8%; Pred. No. 10;
 Matches 11; Conservative 1; Mismatches 1; Indels 3; Gaps 1;

QY 4 NSRGKGRGQGRGNRG 19

Db 242 NSRGGRSQ---NRG 254
 ||||| ||| |||

RESULT 13
 AH1867
 Hypothetical protein alr0489 [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp.
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
 C:Accession: AH1867
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigunakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Nostoc sp. strain PCC 7120
 A:Reference number: ABL807; MUID:21595285; PMID:11759840
 A:Accession: AH1867
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-893 <KUR>
 A:Cross-references: GB:BA000019; PIDN:BAH72447.1; PID:g17129834; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: alr0489
 C:Superfamily: yeast probable SKI2 protein YJL050w

Query Match 47.6%; Score 49.5; DB 2; Length 893;
 Best Local Similarity 61.1%; Pred. No. 28;
 Matches 11; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 1 NPE-NSRGKGRGQGRGNRG 17
 ||| ||| ||| |||
 Db 224 NPRLNRGKRGQGRGNRG 241

RESULT 14
 T39586
 rna binding protein - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T39586
 R:Volckaert, G.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
 A:Submitted to the EMBL Data Library, August 1997
 A:Reference number: Z21865
 A:Accession: T39586
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-166 <VOL>
 A:Cross-references: EMBL:Z99759; PIDN:CAB16904.1; GSPDB:GN00067; SPDB:SPBCL6E9.12c
 A:Experimental source: strain 972h-; cosmid c16E9
 C:Genetics:
 A:Gene: SPDB:SPBCL6E9.12c
 A:Map position: 2
 A:Introns: 12/3; 97/2; 126/3

Query Match 47.1%; Score 49; DB 2; Length 166;
 Best Local Similarity 55.6%; Pred. No. 7.3;
 Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 PENSCKGRRGQGRGNRG 19
 | ||| ||| |||
 Db 135 PGMSRGRRGGRGGRG 152

RESULT 15
 S36541
 E2 protein - human papillomavirus type 12
 C:Species: human papillomavirus type 12
 C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
 C:Accession: S36541
 R:Deilus, H.; Hofmann, B.
 A:Submitted to the EMBL Data Library, August 1993
 A:Description: Primer-directed sequencing of human papillomavirus types.

A:Reference number: S36469
A:Accession: S36541
A:Molecule type: DNA
A:Residues: 1-494
A:Cross-references: EMBL:X74466; NID:g396910; PIDN:CAA52499.1; PID:g396914
C:Superfamily: papillomavirus E2 protein
C:Keywords: DNA binding; early protein; transcription regulation

Query Match	47.1%	Score 49;	DB 2;	Length 494;
Best Local Similarity	69.2%	Pred. No. 19;		
Matches	9;	Conservative	1;	Mismatches 3; Indels 0; Gaps 0;

OY 6 RKGRRGQGGKNR 18
||:|||||
Db 344 RGRGRRGGRKR 356

Search completed: December 4, 2002, 08:19:16
Job time : 18 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 08:15:32 ; Search time 10 Seconds
(without alignments)
78.805 Million cell updates/sec

Title: US-09-687-993-18
Perfect score: 104
Sequence: 1 NPENSRGKGRGQKNGK 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

atched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	104	100.0	211	1 GDNF_HUMAN	P39905 homo sapien
2	99	95.2	211	1 GDNF_MOUSE	P48540 mus musculus
3	99	95.2	211	1 GDNF_RAT	Q07731 rattus norv
4	94	51.9	708	1 MR11_HUMAN	P49959 homo sapien
5	51.5	49.5	894	1 ILF3_HUMAN	Q12906 h interleuk
6	51.5	49.5	910	1 ILF3_RAT	Q9j1l3 rattus norv
7	51.5	49.5	911	1 ILF3_MOUSE	Q921x4 mus musculus
8	51	49.0	525	1 NAB2_YEAST	P32505 saccharomyc
9	50	48.1	265	1 RS2_LETAM	O43992 leishmania
10	50	48.1	340	1 LRP1_HSV1F	P17588 herpes simp
11	50	48.1	613	1 DEAD_HAEIN	P44586 haemophilus
12	49	47.1	494	1 VE2_HPV12	P36782 human papil
13	48.5	46.6	633	1 ROR_HUMAN	O43390 homo sapien
14	48.5	46.6	643	1 YK09_CAEEL	P34304 caenorhabdi
15	48	46.2	126	1 SMD1_CAEEL	Q10013 caenorhabdi
16	48	46.2	758	1 PARC_RHIME	Q59749 rhizobium m
17	47.5	45.7	117	1 SMD1_SCHPO	O42661 schizosacch
18	47	45.2	706	1 MR11_MOUSE	Q61216 mus musculus
19	46	44.2	65	1 Y214_TREPA	O83244 treponema p
20	46	44.2	321	1 FBRL_HUMAN	P22087 homo sapien
21	46	44.2	551	1 CC14_YEAST	Q00684 saccharomyc
22	46	44.2	632	1 FMRI_HUMAN	Q06787 homo sapien
23	46	44.2	706	1 MR11_RAT	Q9j1m0 rattus norv
24	45.5	43.8	123	1 LSM4_CAEEL	Q19952 caenorhabdi
25	45	43.3	178	1 VP12_WTVNJ	P31612 wound tumor
26	45	43.3	253	1 RS2_SCHPO	O74892 schizosacch
27	45	43.3	707	1 RHO_STRLI	P52157 streptomyc
28	45	43.3	862	1 TP3B_MOUSE	Q9z321 mus musculus
29	44.5	42.8	153	1 IF1A_YEAST	P38912 saccharomyc
30	44.5	42.8	414	1 NSR1_YEAST	P27476 saccharomyc
31	44.5	42.8	700	1 MR11_CHICK	Q9iam7 gallus gall
32	44	42.3	483	1 VE2_HPV14	P36783 human papil
33	44	42.3	720	1 MR11_ARATH	Q9xgm2 arabidopsis

RESULT 1				
GDNF_HUMAN				
ID	GDNF_HUMAN	STANDARD;	PRT;	211 AA.
AC	P39905; Q9UP97; Q9UD33;			
DT	01-FEB-1995 (Rel. 31, Created)			
DT	01-FEB-1995 (Rel. 31, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Glial cell line-derived neurotrophic factor precursor (Astrocyte-			
DE	derived trophic factor 1) (ATF-1).			
GN	GDNF.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID:9606;			
RN	[1]			
RN	SEQUENCE FROM N.A. (ISOFORM 1).			
RX	MEDLINE-93262463; PubMed-8493557;			
RA	Lin L.F.H., Doherty D.H., Lille J.D., Bektesh S., Collins F.;			
RT	"GDNF: a glial cell line-derived neurotrophic factor for midbrain			
RT	dopaminergic neurons.";			
RL	Science 260:1130-1132(1993).			
RN	[2]			
RN	SEQUENCE FROM N.A. (ISOFORM 2).			
RX	MEDLINE-95172201; PubMed-7867768;			
RA	Schaar D.G., Sieber B.A., Sherwood A.C., Dean D., Mendoza G.,			
RA	Ramakrishnan L., Dreyfus C.F., Black I.B.;			
RT	"Multiple astrocyte transcripts encode nigral trophic factors in			
RT	and human.";			
RL	Exp. Neurol. 130:387-393(1994).			
RN	[3]			
RP	SEQUENCE OF 1-187 FROM N.A. (ISOFORM 1).			
RC	TISSUE-Kidney;			
RX	MEDLINE-99296655; PubMed-10366742;			
RA	Baecker P.A., Lee W.H., Verity A.N., Eglen R.M., Johnson R.M.;			
RT	"Characterization of a promoter for the human glial cell line-de-			
RT	neurotrophic factor gene.";			
RL	Brain Res. Mol. Brain Res. 69:209-222(1999).			
RN	[4]			
RP	PARTIAL SEQUENCE, AND DISULFIDE BONDS.			
RX	MEDLINE-97141760; PubMed-8988018;			
RA	Haniu M., Hui J., Young Y., Le J., Katta V., Lee R., Shimamoto G.;			
RA	Rohde M.F.;			
RT	"Glial cell line-derived neurotrophic factor: selective reduction			
RT	of the intermolecular disulfide linkage and characterization of its			
RT	disulfide structure.";			
RL	Biochemistry 35:16799-16805(1996).			
RN	[5]			
RP	REVIEW ON VARIANTS.			
RP	MEDLINE-98023959; PubMed-9359036;			
RX	Hofstra R.M.W., Oasinga J., Buys C.H.C.M.;			
RA	"Mutations in Hirschsprung disease: when does a mutation contribu-			
RT	te to the phenotype.";			
RL	Eur. J. Hum. Genet. 5:180-185(1997).			
RN	[6]			
RP	VARIANT HSCR SER-154.			
RX	MEDLINE-97123511; PubMed-8968758;			

Q15027 homo sapien
P22082 saccharomyc
Q09003 xenopus lae
P13941 rattus norv
P08121 mus musculus
P34486 caenorhabdi
Q27294 drosophila
O9J125 neisseria m
O9K012 neisseria m
P12978 epstein-bar
P26545 human papil
P27565 sendai viru

RA Ivanchuk S.M., Myers S.M., Eng C., Mulligan L.M.;
 RT "De novo mutation of GDNF, ligand for the RET/GDNFR-alpha receptor
 complex, in Hirschsprung disease."; Hum. Mol. Genet. 5:2023-2026(1996).
 RL [7]
 RN VARIANT HSCR TRP-93.
 RX MEDLINE=97051933; PubMed=8896568;
 RA Angrist M., Bolk S., Halushka M., Lapchak P.A., Chakravarti A.;
 RT "Germline mutations in glial cell line-derived neurotrophic factor
 (GDNF) and RET in a Hirschsprung disease patient."; Nat. Genet. 14:341-344(1996).
 RL [8]
 RN VARIANTS HSCR SER-21 AND ASN-150.
 RX MEDLINE=97051934; PubMed=8896569;
 RA Salomon R., Attie T., Pelet A., Bidaud C., Eng C., Amiel J.,
 RA Sarnacki S., Goulet O., Ricour C., Nihoul-Fekete C., Munnich A.,
 RA Lyonnet S.;
 RT "Germline mutations of the RET ligand GDNF are not sufficient to cause
 Hirschsprung disease."; Nat. Genet. 14:345-347(1996).
 RL [9]
 CC -1- FUNCTION: NEUROTROPHIC FACTOR THAT ENHANCES SURVIVAL AND
 MORPHOLOGICAL DIFFERENTIATION OF DOPAMINERGIC NEURONS AND
 INCREASES THEIR HIGH-AFFINITY DOPAMINE UPTAKE.
 CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- DISEASE: IN ASSOCIATION WITH MUTATIONS OF RET GENE, MAY BE
 CC INVOLVED IN HIRSCHSPRUNG'S DISEASE (HSCR). THIS GENETIC DISORDER
 CC OF NEURAL CREST DEVELOPMENT IS CHARACTERIZED BY THE ABSENCE OF
 CC INTRAMURAL GANGLION CELLS IN THE HINDGUT; OFTEN RESULTING IN
 CC INTESTINAL OBSTRUCTION.
 CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.
 CC -1- DATABASE: NAME=RD Systems' cytokine mini-reviews: GDNF;
 CC WWW="http://www.rndsystems.com/asp/g_sitebuilder.asp?bodyid=201".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L19063; AAA67910.1;
 DR EMBL: L19062; AAA67910.1; JOINED.
 DR EMBL: AF053748; AAD43139.1;
 DR PIR: B37499; B37499.
 DR HSP: Q07731; IAGO.
 DR Genew: HGNC:4232; GDNF.
 DR MIM: 600837;
 DR MIM: 142623;
 DR InterPro: IPR001839; TGFb.
 DR Pfam: PF00019; TGF-beta; 1.
 DR SMART: SM00204; TGFb; 1.
 DR PROSITE: PS00250; TGF-BETA_1; FALSE_NEG.
 DR KW Growth factor; Glycoprotein; Signal; Alternative splicing;
 KW Polymorphism; Disease mutation; Hirschsprung disease.
 FT SIGNAL 1 19 POTENTIAL.
 FT PROPP 20 77 BY SIMILARITY.
 FT CHAIN 78 211 GLIAL CELL LINE-DERIVED NEUROTROPHIC
 FT FACTOR.
 FT FT DISULFID 118 179
 FT FT DISULFID 145 208
 FT FT DISULFID 149 210
 FT FT DISULFID 178 210
 FT FT CARBOHYD 126 126
 FT FT CARBOHYD 162 162
 FT FT VARSPLIC 25 51
 FT FT VARIANT 21 21
 FT FT P -> S (IN HSCR; COULD BE A
 FT FT POLYMORPHISM).
 FT FT /FTID-VAR_009494.

FT VARIANT 93 93 R -> W (IN HSCR; ASSOCIATED TO A RET
 FT MUTATION; COULD BE AN EXTREMELY RARE
 FT POLYMORPHISM).
 FT /FTID-VAR_009495.
 FT D -> N (IN HSCR; COULD BE A
 FT POLYMORPHISM).
 FT /FTID-VAR_009496.
 FT T -> S (IN HSCR; SPORADIC FORM).
 FT /FTID-VAR_009497.
 FT SEQUENCE 211 AA; 23720 MW; A0D1EBF77FC82691 CRC64;
 SQ
 Query Match 100.0%; Score 104; DB 1; Length 211;
 Best Local Similarity 100.0%; Pred. No. 4e-08;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NPENSRGKGRGGRGKNGR 19
 Db 99 NPENSRGKGRGGRGKNGR 117
 |||||
 RESULT 2
 GDNF_MOUSE STANDARD; PRT; 211 AA.
 ID AC P48540; P97919; O09058; P97920; P70446;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glial cell line-derived neurotrophic factor precursor.
 GN GDNF.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RC STRAIN=C57BL/10J; TISSUE=Brain;
 RA Wang F., Too H.P.;
 RN Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ICR; TISSUE=Dorsal root ganglion;
 RX MEDLINE=95379105; PubMed=7650763;
 RA Watabe K., Fukuda T., Tanaka J., Honda H., Toyohara K., Sakai O.;
 RT "Spontaneously immortalized adult mouse Schwann cells secrete
 RT autocrine and paracrine growth-promoting activities."; J.
 RN J. Neurosci. Res. 41:279-290(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SvJ;
 RA Hellmich H., Kos L., Cho E.S., Mahon K.A., Zimmer A.;
 RN Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Matsushita N., Fujita Y., Nagatsu T., Kiuchi K.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: NEUROTROPHIC FACTOR THAT ENHANCES SURVIVAL AND
 MORPHOLOGICAL DIFFERENTIATION OF DOPAMINERGIC NEURONS AND
 INCREASES THEIR HIGH-AFFINITY DOPAMINE UPTAKE.
 CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U37459; AAB18672.1; ALT_INIT.

```
DR EMBL; U66195; AAB07463.1; ALT_INIT.
DR EMBL; U75532; AAB18343.1; ALT_INIT.
DR EMBL; D49921; BAA08660.1; -.
DR EMBL; U36449; AAB2953.1; -.
DR EMBL; D88264; BAA13566.1; ALT_INIT.
DR EMBL; D88352; BAB12221.1; -.
DR EMBL; D88351; BAB12221.1; JOINED.
DR HSSP; Q07731; IAGO.
DR MGD; MGI:107430; Gdnf.
DR InterPro; IPR001839; TGFb.
DR Pfam; PF00019; TGF-beta; 1.
DR SMART; SM00204; TGFb; 1.
DR PROSITE; PS00250; TGF_BETA_1; FALSE_NEG.
KW Growth factor; Glycoprotein; Signal; Alternative splicing.
FT SIGNAL 1 19
FT PROPEP 20 77
FT CHAIN 78 211
FT DISULFID 118 179
FT DISULFID 145 208
FT DISULFID 149 210
FT DISULFID 178 178
FT CARBOHYD 126 126
FT CARBOHYD 162 162
FT VARSPPLIC 25 51
SQ SEQUENCE 211 AA; 23662 MW; B6731C767A3A95B7 CRC64;

Query Match 95.2%; Score 99; DB 1; Length 211;
Best Local Similarity 94.7%; Pred. No. 2.1e-07;
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NPENSRGKGRGGRGKNGR 19
Db :|||||
99 SPENSRGKGRGGRGKNGR 117

RESULT 3
GDNF_RAT
ID GDNF_RAT STANDARD; PRT; 211 AA.
AC Q07731; Q64062; Q64063; Q63214;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glial cell line-derived neurotrophic factor precursor.
GN GDNF.
Rattus norvegicus (Rat).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A. (ISOFORM 1), AND SEQUENCE OF 78-102.
RX MEDLINE-93262463; PubMed-8493557;
RA Lin L.-F.H., Doherty D.H., Lille J.D., Bektess S., Collins F.;
*GDNF: a glial cell line-derived neurotrophic factor for midbrain
RT dopaminergic neurons.*
RL Science 260:1130-1132(1993).
RN [2]
SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX MEDLINE-95203379; PubMed-7895811;
RA Springer J.E., Seeburger J.L., He J., Gabrea A., Blankenhorn E.P.,
Bergman L.W.;
*cdna sequence and differential mRNA regulation of two forms of glial
RT cell line-derived neurotrophic factor in Schwann cells and rat
skeletal muscle.*
RL Exp. Neurol. 131:47-52(1995).
RN [3]
SEQUENCE OF 1-50 FROM N.A., AND ALTERNATIVE SPLICING.
RP STRAIN-Wistar; TISSUE-Kidney;
RX MEDLINE-95210610; PubMed-7656586;
RA Suter-Crazzolara C., Unsicker K.;
*GDNF is expressed in two forms in many tissues outside the CNS.*;
RL NeuroReport 5:2486-2488(1994).
```

```
RN [4]
RX ALTERNATIVE SPLICING.
RX MEDLINE-95172201; PubMed-7867768;
RA Schaar D.G., Sieber B.A., Sherwood A.C., Dean D., Mendoza G.,
Ramakrishnan L., Dreyfus C.F., Black I.B.;
*Multiple astrocyte transcripts encode nigral trophic factors in rat
RT and human.*;
RL Exp. Neurol. 130:387-393(1994).
RN [5]
RX X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE-97331316; PubMed-9187648;
RA Eigenbrot C., Gerber C.;
*X-ray structure of glial cell-derived neurotrophic factor at 1.9-A
RT resolution and implications for receptor binding.*;
RL Nat. Struct. Biol. 4:435-438(1997).
CC -1- FUNCTION: NEUROTROPHIC FACTOR THAT ENHANCES SURVIVAL AND
MORPHOLOGICAL DIFFERENTIATION OF DOPAMINERGIC NEURONS AND
INCREASES THEIR HIGH-AFFINITY DOPAMINE UPTAKE.
CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1/GDNF633 (SHOWN HERE) AND
2/GDNF555; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L15305; AAB67909.1; -.
DR EMBL; S75583; AAB33891.1; -.
DR EMBL; S75585; AAB33892.1; -.
DR EMBL; X92495; CAA63237.1; -.
DR PIR; A37499; A37499.
DR PDB; IAGO; 05-JUN-97.
DR InterPro; IPR001839; TGFb.
DR Pfam; PF00019; TGF-beta; 1.
DR SMART; SM00204; TGFb; 1.
DR PROSITE; PS00250; TGF_BETA_1; FALSE_NEG.
KW Growth factor; Glycoprotein; Signal; Alternative splicing;
KW 3D-structure.
FT SIGNAL 1 19
FT PROPEP 20 77
FT CHAIN 78 211
FT DISULFID 118 179
FT DISULFID 145 208
FT DISULFID 149 210
FT DISULFID 178 178
FT CARBOHYD 126 126
FT CARBOHYD 162 162
FT VARSPPLIC 25 51
FT CONFLICT 77 77
FT CONFLICT 90 90
FT CONFLICT 101 101
SQ SEQUENCE 211 AA; 23619 MW; AE06C646682895A5 CRC64;

Query Match 95.2%; Score 99; DB 1; Length 211;
Best Local Similarity 94.7%; Pred. No. 2.1e-07;
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NPENSRGKGRGGRGKNGR 19
Db :|||||
99 SPENSRGKGRGGRGKNGR 117

RESULT 4
MR11_HUMAN
ID MR11_HUMAN STANDARD; PRT; 708 AA.
```

P49959; 043475;
 01-OCT-1996 (Rel. 34, Created)
 16-OCT-2001 (Rel. 40, Last sequence update)
 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Double-strand break repair protein MRE11A (MRE11 homolog 1).
 GN MRE11A OR MRE11 OR HNGS1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 [1]
 SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=96079094; PubMed=8530104;
 RA Petrini J.H.J., Walsh M.E., Dimare C., Chen X.-N., Korenberg J.R.,
 RA Weaver D.T.;
 RT "Isolation and characterization of the human MRE11 homologue.";
 RL Genomics 29:80-86(1995).
 [2]
 REVISIONS TO C-TERMINUS.
 RA Petrini J.H.J., Walsh M.E., Dimare C., Chen X.-N., Korenberg J.R.,
 RA Weaver D.T.;
 RT Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 [3]
 SEQUENCE FROM N.A. (ISOFORM 1).
 RA Chamankhah M., Wei Y., Xiao W.;
 RT "Molecular cloning and functional characterization of hNGS1, a yeast
 and human MRE11 homolog.";
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 [4]
 SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=98315380; PubMed=9651580;
 RA Pauli T.T., Gellert M.;
 RT "The 3' to 5' exonuclease activity of Mre 11 facilitates repair of DNA
 double-strand breaks.";
 RL Mol. Cell 1:969-979(1998).
 [5]
 SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=21264328; PubMed=11371508;
 RA Pitts S.A., Kullar H.S., Stankovic T., Stewart G.S., Last J.I.K.,
 RA Bedenham T., Armstrong S.J., Plane M., Chessa L., Taylor A.M.R.,
 RA Byrd P.J.;
 RT "hMRE11: genomic structure and a null mutation identified in a
 transcript protected from nonsense-mediated mRNA decay.";
 RL Hum. Mol. Genet. 10:1155-1162(2001).
 [6]
 VARIANT ATLD SER-117.
 RX MEDLINE=20077641; PubMed=10612394;
 RA Stewart G.S., Maser R.S., Stankovic T., Bressan D.A., Kaplan M.I.,
 RA Jaspers N.G.J., Raams A., Byrd P.J., Petrini J.H.J., Taylor A.M.R.;
 RT "The DNA double-strand break repair gene hMRE11 is mutated in
 individuals with an ataxia-telangiectasia-like disorder.";
 RL Cell 99:577-587(1999).
 [7]
 VARIANTS.
 RX MEDLINE=21036682; PubMed=11196167;
 RA Fukuda T., Sumiyoshi T., Takahashi M., Kataoka T., Asahara T.,
 RA Inui H., Watanabe M., Yasutomi M., Kanada N., Miyagawa K.;
 RT "Alterations of the double-strand break repair gene MRE11 in cancer.";
 RL Cancer Res. 61:23-26(2001).
 CC -!- FUNCTION: Involved in DNA double-strand break repair (DSBR).
 CC Possesses single-strand endonuclease activity and double-strand-
 CC specific 3'-5' exonuclease activity. Also involved in meiotic DSB
 CC processing.
 CC -!- COFACTOR: Manganese (By similarity).
 CC -!- SUBUNIT: FORMS A COMPLEX WITH RAD50 (By similarity).
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
 CC produced by alternative splicing.
 CC -!- DISEASE: DEFECTS IN MRE11A ARE A CAUSE OF ATAXIA-TELANGIECTASIA-
 CC LIKE DISORDER (ATLD). IT IS A DISEASE WITH THE SAME CLINICAL
 CC FEATURE THAN ATAXIA-TELANGIECTASIA BUT WITH A SOMEWHAT Milder
 CC CLINICAL COURSE.
 CC -!- SIMILARITY: BELONGS TO THE MRE11/RAD32 FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL: U37359; AAC78721.1; -
 EMBL: AF022778; AAD10197.1; -
 EMBL: AF073362; AAC36249.1; -
 EMBL: AF303395; AAK18790.1; -
 EMBL: AF303379; AAK18790.1; JOINED.
 EMBL: AF303380; AAK18790.1; JOINED.
 EMBL: AF303381; AAK18790.1; JOINED.
 EMBL: AF303382; AAK18790.1; JOINED.
 EMBL: AF303383; AAK18790.1; JOINED.
 EMBL: AF303384; AAK18790.1; JOINED.
 EMBL: AF303385; AAK18790.1; JOINED.
 EMBL: AF303386; AAK18790.1; JOINED.
 EMBL: AF303387; AAK18790.1; JOINED.
 EMBL: AF303388; AAK18790.1; JOINED.
 EMBL: AF303389; AAK18790.1; JOINED.
 EMBL: AF303390; AAK18790.1; JOINED.
 EMBL: AF303391; AAK18790.1; JOINED.
 EMBL: AF303392; AAK18790.1; JOINED.
 EMBL: AF303393; AAK18790.1; JOINED.
 EMBL: AF303394; AAK18790.1; JOINED.
 Genew: HGNC:7230; MRE11A.
 MIM: 600814; -
 MIM: 604391; -
 InterPro: IPR003701; DNA_repair.
 InterPro: IPR004843; M-ppetrase.
 InterPro: IPR004844; S/T-phosphatase.
 Pfam: PF00149; Metallophos; 1.
 TIGRFAMs: TIGR00583; mre11; 1.
 DNA repair: Hydrolase; Nuclease; Endonuclease; Exonuclease;
 Nuclear protein; Manganese; Meiosis; Alternative splicing;
 Disease mutation; Polymorphism.
 VARSPLIC 595 622 MISSING (IN ISOFORM 2).
 VARIANT 104 104 S -> C (IN CANCER).
 VARIANT 117 117 N -> S (IN ATLD).
 VARIANT 157 157 /FTID-VAR_008513.
 VARIANT 157 157 /FTID-VAR_011626.
 VARIANT 503 503 R -> H (IN CANCER).
 VARIANT 572 572 /FTID-VAR_011627.
 VARIANT 572 572 R -> Q (IN CANCER).
 CONFLICT 31 31 /FTID-VAR_011628.
 CONFLICT 31 31 V -> A (IN REF.1).
 SQ SEQUENCE 708 AA; 80593 MW; D94ABFDDDF6106AD CRC64;
 Query Match 51.9%; Score 54; DB 1; Length 708;
 Best Local Similarity 58.8%; Pred. No. 1.8;
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 QY 1 NPENSRGKGRGRGQGN 17
 DB 567 NKGRGRGRGRGRGQGN 583
 RESULT 5
 ILF3_HUMAN
 ID ILF3_HUMAN STANDARD; PRT; 894 AA.
 AC Q12906; Q9UN84; Q9NQ96; Q9NQ97; Q9NQ98; Q9NQ99; Q9NQAO;
 AC Q9NOA1; Q9NOA2; Q9UN29; Q9UN29; Q9UN00; Q9B2H4; Q9B2H5; Q99544;
 AC Q99545; Q9NRN2; Q9NRN3; Q9NRN4; O43409;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Interleukin enhancer-binding factor 3 (Nuclear factor of activated T

DE cells-90) (NF-AT-90) (Double-stranded RNA-binding protein 76) (DRBP76)
DE (Translational control protein 80) (TCP80) (Nuclear factor associated
DE with GARN) (NFAR) (M-phase phosphoprotein 4) (MPP4).
GN ILF3 OR NF90 OR DRBF OR MPOSPH4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP PRELIMINARY SEQUENCE FROM N.A., AND SEQUENCE OF 19-41; 491-510 AND
RP 555-565.
RC TISSUE=T-cell lymphoma;
RX MEDLINE=94327652; PubMed=7519613;
RA Kao P.N., Chen L., Brock G., Ng J., Kenny J., Smith A.J., Cortesby B.;
RT "Cloning and expression of cyclosporin A- and FK506-sensitive nuclear
RT factor of activated T-cells: NF45 and NF90.";
RL J. Biol. Chem. 269:20691-20699(1994).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2), AND SEQUENCE OF N-TERMINUS.
RP TISSUE=Cervical carcinoma;
RX MEDLINE=99329057; PubMed=10400669;
RA Patel R.C., Vestal D.J., Xu Z., Bandyopadhyay S., Guo W., Erme S.M.,
RA Williams B.R., Sen G.C.;
RT "DRBP76, a double-stranded RNA-binding nuclear protein, is
RT phosphorylated by the interferon-induced protein kinase, PKR.";
RL J. Biol. Chem. 274:20432-20437(1999).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 3), AND CHARACTERIZATION.
RP TISSUE=Liver;
RX MEDLINE=20076798; PubMed=10607473;
RA Xu Y.-H., Grabowski G.A.;
RT "Molecular cloning and characterization of a translational inhibitory
RT protein that binds to coding sequences of human acid beta-glucosidase
RT and other mRNAs.";
RL Mol. Genet. Metab. 68:441-454(1999).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 4 AND 5), AND ALTERNATIVE SPLICING.
RP TISSUE=Melanoma;
RX MEDLINE=21100430; PubMed=11167023;
RA Duchange N., Pidoux J., Canus E., Sauvaget D.;
RT "Alternative splicing in the human interleukin enhancer binding factor
RT 3 (ILF3) gene.";
RL Gene 261:345-353(2000).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND CHARACTERIZATION.
RP MEDLINE=21402983; PubMed=11438536;
RX Saunders L.R., Perkins D.J., Balachandran S., Michaels R., Ford R.,
RX Mayeda A., Barber G.N.;
RT "Characterization of two evolutionarily conserved, alternatively
RT spliced nuclear phosphoproteins, NFAR-1 and -2, that function in mRNA
RT processing and interact with the double-stranded RNA-dependent
RT protein kinase, PKR.";
RL J. Biol. Chem. 276:32300-32312(2001).
RN [6]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RP TISSUE=T-cell;
RX PubMed=1161820;
RA Saunders L.R., Juretic V., Barber G.N.;
RT "The 90- and 110-kDa human NFAR proteins are translated from two
RT differentially spliced mRNAs encoded on chromosome 19p13.";
RL Genomics 71:256-259(2001).
RN [7]
RP SEQUENCE OF 1-611 FROM N.A.
RP TISSUE=Cervix, and Blood;
RX MEDLINE=97039687; PubMed=8885239;
RA Matsumoto-Taniguchi N., Pirollet F., Monroe R., Gerace L.,
RA Westendorp J.M.;
RT "Identification of novel M phase phosphoproteins by expression
RT cloning.";
RL Mol. Biol. Cell 7:1455-1469(1996).
RN [8]
RP SEQUENCE OF 188-894 FROM N.A. (ISOFORMS 1; 2 AND 3).
RP MacArdle J., Cantarella G.M., Veyrune J.-L., Krasnoselskaya I.,
Kumar A.;
RT "Structure and functional characterization of hDRBF gene.";
RT Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [9]
RP SEQUENCE OF 587-894 FROM N.A. (ISOFORM 1).
RP TISSUE=Brain;
RC Yu W., Sarginson J., Gibbs R.A.;
RA Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [10]
RP INTERACTION WITH HR23L2.
RX MEDLINE=20347897; PubMed=10749851;
RA Tang J., Kao P.N., Herschman H.R.;
RT "Protein-arginine methyltransferase I, the predominant
RT protein-arginine methyltransferase in cells, interacts with and is
RT regulated by interleukin enhancer-binding factor 3.";
RL J. Biol. Chem. 275:19866-19876(2000).
CC -!- FUNCTION: May facilitate double-stranded RNA-regulated gene
CC expression at the level of post-transcription. Can act as a
CC translation inhibitory protein which binds to coding sequences of
CC acid beta-glucosidase (GCase) and other mRNAs and functions at the
CC initiation phase of GCase mRNA translation, probably by inhibiting
CC its binding to polysomes. Can regulate protein arginine N-
CC methyltransferase 1 activity.
CC -!- SUBUNIT: Interacts with FUS and SMN proteins and also with
CC HR23L2.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS: 5 isoforms; 1/NFAR-2/ILF3-E (shown here),
CC 2/NFAR-1/DRBP76, 3, 4/DRBP76 Alpha/ILF3-A and 5/DRBP76
CC Delta/Gamma/ILF3-C; are produced by alternative splicing.
CC -!- TISSUE SPECIFICITY: Ubiquitous.
CC -!- PTM: Phosphorylated by RNA-dependent protein kinase (PKR).
CC -!- PTM: Methylated by protein arginine N-methyltransferase 1.
CC -!- SIMILARITY: CONTAINS 2 DRBM (DOUBLE-STRANDED RNA-BINDING) DOMAINS.
CC -!- CAUTION: Ref.1 sequence differs from that shown due to sequencing
CC errors.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; U10324; AAA20994.1; ALT_SEQ.
DR EMBL; AF147209; RAD33966.1;
DR EMBL; AF141870; RAD37575.1;
DR EMBL; AJ271743; CAC01121.1;
DR EMBL; AJ271741; CAC01122.1;
DR EMBL; AJ271741; CAC01123.1;
DR EMBL; AJ271741; CAC01124.1;
DR EMBL; AJ271744; CAC01404.1;
DR EMBL; AJ271745; CAC01405.1;
DR EMBL; AJ271746; CAC01406.1;
DR EMBL; AJ271747; CAC01407.1;
DR EMBL; AF167569; RAD51098.1;
DR EMBL; AF167570; RAD51099.1;
DR EMBL; AF320244; AAK07424.1;
DR EMBL; AF320228; AAK07424.1; JOINED.
DR EMBL; AF320229; AAK07424.1; JOINED.
DR EMBL; AF320230; AAK07424.1; JOINED.
DR EMBL; AF320231; AAK07424.1; JOINED.
DR EMBL; AF320232; AAK07424.1; JOINED.
DR EMBL; AF320233; AAK07424.1; JOINED.
DR EMBL; AF320234; AAK07424.1; JOINED.
DR EMBL; AF320235; AAK07424.1; JOINED.
DR EMBL; AF320236; AAK07424.1; JOINED.
DR EMBL; AF320237; AAK07424.1; JOINED.
DR EMBL; AF320238; AAK07424.1; JOINED.
DR EMBL; AF320239; AAK07424.1; JOINED.
DR EMBL; AF320240; AAK07424.1; JOINED.
DR EMBL; AF320241; AAK07424.1; JOINED.
DR EMBL; AF320242; AAK07424.1; JOINED.

```
DR EMBL; AF320243; AAK07424.1; JOINED.
DR EMBL; AF320247; AAK07425.1; -.
DR EMBL; AF320228; AAK07425.1; JOINED.
DR EMBL; AF320229; AAK07425.1; JOINED.
DR EMBL; AF320230; AAK07425.1; JOINED.
DR EMBL; AF320231; AAK07425.1; JOINED.
DR EMBL; AF320232; AAK07425.1; JOINED.
DR EMBL; AF320233; AAK07425.1; JOINED.
DR EMBL; AF320234; AAK07425.1; JOINED.
DR EMBL; AF320235; AAK07425.1; JOINED.
DR EMBL; AF320236; AAK07425.1; JOINED.
DR EMBL; AF320237; AAK07425.1; JOINED.
DR EMBL; AF320238; AAK07425.1; JOINED.
DR EMBL; AF320239; AAK07425.1; JOINED.
DR EMBL; AF320240; AAK07425.1; JOINED.
DR EMBL; AF320241; AAK07425.1; JOINED.
DR EMBL; AF320242; AAK07425.1; JOINED.
DR EMBL; AF320243; AAK07425.1; JOINED.
DR EMBL; AF320245; AAK07425.1; JOINED.
DR EMBL; AF320246; AAK07425.1; JOINED.
DR EMBL; X98264; CAA66917.1; -.
DR EMBL; X98265; CAA66918.1; -.
DR EMBL; AF202445; AAF82685.1; -.
DR EMBL; AF202445; AAF82686.1; -.
DR EMBL; AF202445; AAF82687.1; -.
DR EMBL; AF007140; AAC19152.1; -.
DR Genew; HGNC:6038; ILF3.
DR MIM; 603182; -.
DR InterPro; IPR001159; DS_RBD.
DR Pfam; PF00035; dsrm; 2.
DR SMART; SM00358; DSRM; 2.
DR PROSITE; PS50137; DS_RBD; 2.
KW Transcription regulation; DNA-binding; RNA-binding; Nuclear protein;
KW Repeat; Phosphorylation; Methylation.
FT DOMAIN 371 389
BIPARTITE NUCLEAR LOCALIZATION SIGNAL
(POTENTIAL).
FT DOMAIN 398 467
DRBM 1.
FT DOMAIN 524 590
DRBM 2.
FT DOMAIN 609 894
INTERACTS WITH HRMTLL2.
FT DOMAIN 640 659
ARG/GLY-RICH.
FT DOMAIN 385 389
POLY-LYS.
FT DOMAIN 634 637
POLY-PRO.
FT DOMAIN 701 709
POLY-GLY.
FT DOMAIN 794 798
POLY-GLY.
FT VARSPIC 516 516
E -> ENVKQ (IN ISOFORM 4).
FT VARSPIC 688 694
QFYSNGG -> KCAFLSV (IN ISOFORM 4).

Query Match 49.5%; Score 51.5; DB 1; Length 894;
Best Local Similarity 57.9%; Pred. No. 5;
Matches 11; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 2 PENSNGKGRGQ-RGKNRG 19
| | | | | | | | | | | | | |
Db 636 PPNLRGGRGGSIRGRGRG 654

RESULT 6
ILF3_RAT
ID ILF3_RAT STANDARD; PRT; 910 AA.
AC Q9JIL3;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Interleukin enhancer-binding factor 3.
GN ILF3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20347897; PubMed=10749851;
RA Tang J., Kao P.N., Herschman H.R.;
```

```
*Protein-arginine methyltransferase 1, the predominant protein-
arginine methyltransferase in cells, interacts with and is regulated
by interleukin enhancer-binding factor 3.*;
J. Biol. Chem. 275:19866-19876(2000).
-!- FUNCTION: May facilitate double-stranded RNA-regulated gene
expression at the level of post-transcription. Can act as a
translation inhibitory protein which binds to coding sequences of
acid beta-glucocidase (GCaase) and other mRNAs and functions at the
initiation phase of GCaase mRNA translation, probably by inhibiting
its binding to polysomes (By similarity). Can regulate protein
arginine N-methyltransferase 1 activity.
-!- SUBUNIT: Interacts with FUS and SMN proteins (By similarity).
Interacts with HRMTLL2.
-!- SUBCELLULAR LOCATION: Nuclear.
-!- PTM: Phosphorylated by RNA-dependent protein kinase (PRKR) (By
similarity).
-!- PTM: Methylated by protein arginine N-methyltransferase 1.
-!- SIMILARITY: CONTAINS 2 DRBM (DOUBLE-STRANDED RNA-BINDING) DOMAINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; AF220102; AAF31446.1; -.
CC InterPro; IPR001159; DS_RBD.
CC Pfam; PF00035; dsrm; 2.
CC SMART; SM00358; DSRM; 2.
CC PROSITE; PS50137; DS_RBD; 2.
KW Transcription regulation; DNA-binding; RNA-binding; Nuclear protein;
KW Repeat; Phosphorylation; Methylation.
FT DOMAIN 384 402
BIPARTITE NUCLEAR LOCALIZATION SIGNAL
(POTENTIAL).
FT DOMAIN 411 480
DRBM 1.
FT DOMAIN 537 603
DRBM 2.
FT DOMAIN 622 910
INTERACTS WITH HRMTLL2.
FT DOMAIN 653 672
ARG/GLY-RICH.
FT DOMAIN 11 15
POLY-ARG.
FT DOMAIN 398 402
POLY-LYS.
FT DOMAIN 647 650
POLY-PRO.
FT DOMAIN 714 723
POLY-GLY.
FT DOMAIN 809 813
POLY-GLY.
SQ SEQUENCE 910 AA; 97679 MW; D6AB97A4B52E442A CRC64;

Query Match 49.5%; Score 51.5; DB 1; Length 910;
Best Local Similarity 57.9%; Pred. No. 5.1;
Matches 11; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 2 PENSNGKGRGQ-RGKNRG 19
| | | | | | | | | | | | | |
Db 649 PPNIRGRGRGNIRGRGRG 667

RESULT 7
ILF3_MOUSE
ID ILF3_MOUSE STANDARD; PRT; 911 AA.
AC Q921X4;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Interleukin enhancer-binding factor 3.
GN ILF3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX TISSUE=Testis;
RX MEDLINE=99269911; PubMed=10337617;
```

NCBI_TaxID=4932;
[1]
SEQUENCE FROM N.A.
STRAIN-BJ926;
MEDLINE=93233636; PubMed=8474438;
Anderson J.T., Wilson S.M., Datar K.V., Swanson M.S.;
*NAB2: a yeast nuclear polyadenylated RNA-binding protein essential
for cell viability.*;
Mol. Cell. Biol. 13:2730-2741(1993).
[2]
SEQUENCE FROM N.A.
Laquin G.;
Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
-!- FUNCTION: THIS ESSENTIAL PROTEIN BINDS TO POLYADENYLATED RNA AND
SINGLE-STRANDED DNA. IT MAY BE INVOLVED NOT ONLY IN RNA PROCESSING
BUT ALSO IN TRANSCRIPTION REGULATION. BELIEVED TO ASSOCIATE
DIRECTLY WITH NASCENT RNA POLYMERASE II TRANSCRIPTS AND REMAIN
ASSOCIATED DURING SUBSEQUENT NUCLEAR RNA PROCESSING REACTIONS.
-!- SUBCELLULAR LOCATION: Nuclear.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).

EMBL: L10288; AAA34819.1; -
EMBL: L08079; AAA34820.1; -
EMBL: 272644; CAA96830.1; -
PIR: B48058; B48058.
SGD: S0003090; NAB2.
Nuclear protein; RNA-binding; Zinc-finger; Repeat.
DOMAIN 97 172 GLN-RICH.
FT DOMAIN 106 116 POLY-GLN.
FT DOMAIN 121 156 10 X 4 AA TANDEM REPEATS OF Q-Q-Q-P.
FT REPEAT 121 124 1.
FT REPEAT 125 128 2.
FT REPEAT 129 132 3.
FT REPEAT 133 136 4.
FT REPEAT 137 140 5.
FT REPEAT 141 144 6.
FT REPEAT 145 148 7.
FT REPEAT 149 152 8.
FT REPEAT 153 156 9.
FT REPEAT 157 160 10 (APPROXIMATE).
FT DOMAIN 209 228 RNA-BINDING RGG-BOX (BY SIMILARITY).
FT ZN_FING 262 278 C3H-TYPE 1.
FT ZN_FING 283 300 C3H-TYPE 2.
FT ZN_FING 340 355 C3H-TYPE 3.
FT ZN_FING 371 386 C3H-TYPE 4.
FT ZN_FING 415 430 C3H-TYPE 5.
FT ZN_FING 437 452 C3H-TYPE 6.
FT ZN_FING 458 473 C3H-TYPE 7.
FT VARIANT 130 157 MISSING (IN YJA512).
SQ SEQUENCE 525 AA; 58321 MW; 40335C3D4658D91 CRC64;

Query Match 49.0%; Score 51; DB 1; Length 525;
Best Local Similarity 76.9%; Pred. No. 3.5;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 7 GKRRGGRGKNGR 19
| | | | | | | | | |
Db 215 GKRRGGRGKNGR 227

RESULT 9
RS2_LEIAM STANDARD; PRT; 265 AA.
ID RS2_LEIAM
AC Q43992;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)

```

FT REPEAT      59      75      2.
SQ SEQUENCE    340 AA; 35604 MW; 2977FA8F887E5451 CRC64;

Query Match      48.1%; Score 50; DB 1; Length 340;
Best Local Similarity 50.0%; Pred. No. 3.2;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 2 PENSRRGKRRGRQKNGK 19
   I  III:|  III:|  I  |
Db 305 PRGSRGRGRGRGRGGG 322

RESULT 11
DEAD_HAEIN STANDARD; PRT; 613 AA.
ID DEAD_HAEIN AC P44586;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cold-shock DEAD-box protein A homolog (ATP-dependent RNA helicase dead
DE homolog).
DN DEAD OR CSDA OR HI0231.
GN Haemophilus influenzae.
OS Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Karlvagn A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Frichman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
CC -1- FUNCTION: HAS A HELIX-STABILIZING ACTIVITY (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U32709; AAC21900.1; -.
CC DR HSSP; Q58083; 1HV8.
CC DR TIGR; HI0231; -.
CC DR InterPro; IPR001410; DEAD.
CC DR InterPro; IPR000629; DEAD_box.
CC DR InterPro; IPR001650; Helicase_C.
CC DR Pfam; PF00270; DEAD; 1.
CC DR Pfam; PF00271; helicase_C; 1.
CC DR SMART; SM00487; HEDXC; 1.
CC DR SMART; SM00490; HELIC; 1.
CC DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
CC KW Hydrolase; Helicase; ATP-binding; RNA-binding;
CC KW Transcription regulation; Complete proteome.
CC NP_BIND 49 56 ATP (POTENTIAL).
CC FT SITE 155 158 DEAD BOX.
CC FT SEQUENCE 613 AA; 63705 MW; 1B826CBDEB1704DF CRC64;

Query Match      48.1%; Score 50; DB 1; Length 613;
Best Local Similarity 44.8%; Pred. No. 5.7;

```

Matches 13; Conservative 3; Mismatches 1; Indels 12; Gaps 2;

QY 3 ENSRG-----KGR-----RQGRGNRG 19
:|||||
Db 556 DNSRGSDDFNGKRGGRGDFRGRGRG 584
:|||||

RESULT 12

VE2_HPV12
ID VE2_HPV12 STANDARD; PRT; 494 AA.
AC P36782;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Regulatory protein E2.
OS E2.
GN Human papillomavirus type 12.

OC Viruses: dsDNA viruses, no RNA stage; Papillomaviridae;
Papillomavirus.
NCBI_TaxID=10604;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=94265501; PubMed=8205838;

RA Delliuss H., Hofmann B.;

RT "Primer-directed sequencing of human papillomavirus types.";

RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).

CC -!- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.

CC IT BINDS TO THE E2RE RESPONSE ELEMENT (5'-ACCCNNNNNGT-3') PRESENT

CC IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER

CC ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING OF E2RE'S POSITION

CC WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS

CC BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION

CC INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA

CC REPLICATION.

CC -!- SUBUNIT: BINDS DNA AS A DIMER.

CC -!- SUBCELLULAR LOCATION: Nuclear.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X74456; CAA52499.1; -

PIR; S36541; S36541.

HSP; P17383; LDHM.

InterPro: IPR000427; E2_C.

DR InterPro: IPR001866; E2_N.

DR Pfam: PF00508; E2_N; 1.

DR Pfam: PF00511; E2_C; 1.

DR ProDom: PD000672; E2_C; 1.

DR ProDom: PD000678; E2_N; 1.

KW Early protein; Transcription regulation; Activator; DNA-binding;

KW Trans-acting factor; DNA replication; Repressor; Nuclear protein.

SQ SEQUENCE 494 AA; 56039 MW; 3C20AFA3560C10EE CRC64;

Query Match 47.1%; Score 49; DB 1; Length 494;

Best Local Similarity 69.2%; Pred. No. 6.4;

Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 RGKGRGRGQGNR 18

:||:|||||

Db 344 RGRGGRGGRKR 356

RESULT 13

ROR_HUMAN

AC ROR_HUMAN STANDARD; PRT; 633 AA.

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Heterogeneous nuclear ribonucleoprotein R (hnRNP R).
GN HNRPR.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98083170; PubMed=9421497;

RA Hassfeld W., Chan E.K.L., Mathison D.A., Portman D., Dreyfuss G.,

RT Steiner G., Tan E.M.;

RT "Molecular definition of heterogeneous nuclear ribonucleoprotein R

RT (hnRNP R) using autoimmune antibody: immunological relationship with

RT hnRNP P.";

RL Nucleic Acids Res. 26:439-445(1998).

CC -!- FUNCTION: COMPONENT OF RIBONUCLEOSOMES, WHICH ARE COMPLEXES OF AT

CC LEAST 20 OTHER DIFFERENT HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEINS

CC (HNRNP). HNRNP PLAY AN IMPORTANT ROLE IN PROCESSING OF PRECURSOR

CC MRNA IN THE NUCLEUS.

CC -!- SUBCELLULAR LOCATION: Nuclear; nucleoplasm.

CC -!- SIMILARITY: CONTAINS 3 RNA RECOGNITION MOTIFS (RRM).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AF000364; AAC39540.1; -

DR HSP; P09651; 1HA1.

DR Genew; HGNC:5047; HNRPR.

DR InterPro: IPR000504; RNA_rec_mot.

DR Pfam: PF00076; Rrm; 3.

DR SMART; SM00360; RRM; 3.

DR PROSITE; PS00303; RRM_RNP_1; 2.

DR PROSITE; PS00303; RRM; 3.

KW Nuclear protein; RNA-binding; Ribonucleoprotein; Repeat.

FT DOMAIN 1 153 ASP/GLU-RICH (ACIDIC).

FT DOMAIN 412 418 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).

FT DOMAIN 165 244 RNA-BINDING (RRM) 1.

FT DOMAIN 246 328 RNA-BINDING (RRM) 2.

FT DOMAIN 341 411 RNA-BINDING (RRM) 3.

FT DOMAIN 447 567 RNA-BINDING (RG-BOX).

FT DOMAIN 462 497 3 X 11 AA APPROXIMATE REPEATS OF D-D-Y-Y-

FT REPEAT 462 471 G-Y-D-Y-H-D-Y.

FT REPEAT 472 482 1 (APPROXIMATE).

FT REPEAT 488 497 3 (APPROXIMATE).

FT DOMAIN 579 633 GLN/ASN-RICH DOMAIN.

SQ SEQUENCE 633 AA; 70943 MW; 088341F645ED46F CRC64;

Query Match 46.6%; Score 48.5; DB 1; Length 633;

Best Local Similarity 58.8%; Pred. No. 9.6;

Matches 10; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 3 ENSRGKGRGCRGKNGR 19

:||:|||||

Db 551 QOQGRGRGSRG-NRG 566

RESULT 14

YKQ9_CAEEL

ID YKQ9_CAEEL STANDARD; PRT; 643 AA.

AC P34304;

DT 01-FEB-1994 (Rel. 28, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Hypothetical protein C06E1.9 in chromosome III.

GN C06E1.9.

OS Caenorhabditis elegans.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 08:16:37 ; Search time 30 Seconds
(without alignments)
130.497 Million cell updates/sec

Title: US-09-687-993-18

Perfect score: 104
Sequence: 1 NPENSRGKRRGQGRGKNGR 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	104	100.0	133	4	Q9UD32
2	104	100.0	185	4	Q96L44
3	101	97.1	160	6	Q97685
4	99	95.2	73	11	Q9QXJ7
5	99	95.2	99	11	Q9QXJ8
6	99	95.2	125	11	Q9QXJ9
7	99	95.2	199	11	Q8R485
8	72	69.2	182	13	Q9IAM2
9	72	69.2	215	13	Q9IAM3
10	71	68.3	143	13	Q8QGE9
11	60	57.7	926	10	Q9LVP1
12	58	55.8	1479	10	Q9ATY5
13	58	55.8	1522	10	Q9LRT2
14	57	54.8	513	10	Q8RZE5
15	55.5	53.4	545	10	Q8S6K4
16	55.5	53.4	2925	12	O89163

ALIGNMENTS

RESULT 1

Q9UD32 PRELIMINARY; PRT; 133 AA.

AC Q9UD32;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE ASTROCYTE-derived TROPHIC factor 2, ATF-2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95172201; PubMed=7867768;
RA Schaar D.G., Sleber B.A., Sherwood A.C., Dean D., Mendoza G.,
RA Ramakrishnan L., Dreyfus C.F., Black I.B.;
RT "Multiple astrocyte transcripts encode nigral trophic factors in rat
and human.";
RL Exp. Neurol. 130:387-393(1994).
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR HSSP: Q07731; IAGO.
DR InterPro: IPR001839; TGFb.
DR Pfam: PF00019; TGF-beta; 1.
DR SMART: SM00204; TGFb; 1.
KW Glycoprotein.

Query Match 100.0%; Score 104; DB 4; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 19; Conservatve 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NPENSRGKRRGQGRGKNGR 19

Db 21 NPENSRGKRRGQGRGKNGR 39

RESULT 2

Q96L44

Q9JGN8 rice grassy
Q9A749 caulobacter
Q94GU5 oryza sativ
Q94IW2 oryza sativ
Q9VWF2 drosophila
Q9IPQ8 cynomolgus
Q9NKA5 neurospora
Q9C2K5 drosophila
Q8UG82 agrobacteri
Q91PK1 arabidopsis
Q9C7Y1 arabidopsis
Q9FTH3 arabidopsis
Q99DC5 human immun
Q99DA5 human immun
Q99DA0 human immun
Q91136 human immun
Q91140 human immun
Q91144 human immun
Q91152 human immun
Q9YNC9 human immun
Q91170 human immun
Q9YNC7 human immun
Q9YNC6 human immun
Q91180 human immun
Q900Y9 human immun
Q99Z38 streptococ
Q51287 neisseria m
Q9SFI4 arabidopsis
Q9MAU9 arabidopsis

ID Q96L44 PRELIMINARY; PRT; 185 AA.
 AC Q96L44;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Glial cell-derived neurotrophic factor isoform.
 GN GDNF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang B., Peng Z., Zhou Y., Peng X., Yuan J., Qiang B.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
 DR EMBL; AY052832; AAL11017.1; -;
 DR InterPro; IPR001839; TGFb.
 DR Pfam; PF00019; TGF-beta; 1.
 KW Glycoprotein.
 KW SEQUENCE 185 AA; 20885 MW; 1988C50DA5EA1B10 CRC64;
 Query Match 100.0%; Score 104; DB 4; Length 185;
 Best Local Similarity 100.0%; Pred. No. 2.2e-08;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NPENSRGKGRGORGKNGR 19
 Db |||||
 Db 73 NPENSRGKGRGORGKNGR 91
 RESULT 3
 O97685 PRELIMINARY; PRT; 160 AA.
 ID O97685
 AC O97685
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Neurotrophic factor (Fragment).
 GN GDNF.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Erjia C., Yong L., QiuJiang D.;
 RL "The gene cloning of macaca and human GDNF by direct PCR from whole
 blood and sequence analysis.";
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
 DR EMBL; AF106678; AAC99782.1; -;
 DR HSSP; Q07731; IAGQ.
 DR InterPro; IPR001839; TGFb.
 DR Pfam; PF00019; TGF-beta; 1.
 DR SMART; SM00204; TGFb; 1.
 KW Glycoprotein.
 FT NON_TER 1
 SQ SEQUENCE 160 AA; 18196 MW; E206362185D499B4 CRC64;
 Query Match 97.1%; Score 101; DB 6; Length 160;
 Best Local Similarity 94.7%; Pred. No. 5.4e-08;
 Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NPENSRGKGRGORGKNGR 19
 Db |||||
 Db 48 NPENSRGKGRGORGKNGR 66
 RESULT 4
 O9QXJ7 PRELIMINARY; PRT; 73 AA.
 ID O9QXJ7

AC Q9QXJ7;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE GDNF splice variant 3 (Fragment).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WISTAR; TISSUE=BRAIN;
 RA Russell F.D., Koishi K., Jiang Y., McLennan I.S.;
 RL "Anterograde axonal transport of glial cell line-derived neurotrophic
 factor and its receptors in rat hypoglossal nerve.";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF205715; AAF23769.1; -;
 DR NON_TER 73
 SQ SEQUENCE 73 AA; 8262 MW; D6394FE64FFC67AB CRC64;
 Query Match 95.2%; Score 99; DB 11; Length 73;
 Best Local Similarity 94.7%; Pred. No. 4.7e-08;
 Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NPENSRGKGRGORGKNGR 19
 Db |||||
 Db 47 SPENSRGKGRGORGKNGR 65
 RESULT 5
 O9QXJ8 PRELIMINARY; PRT; 99 AA.
 ID O9QXJ8
 AC O9QXJ8;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE GDNF splice variant 2 (Fragment).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WISTAR; TISSUE=BRAIN;
 RA Russell F.D., Koishi K., Jiang Y., McLennan I.S.;
 RL "Anterograde axonal transport of glial cell line-derived neurotrophic
 factor and its receptors in rat hypoglossal nerve.";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF205714; AAF23768.1; -;
 DR NON_TER 99
 SQ SEQUENCE 99 AA; 11025 MW; C60C998CD9C58723 CRC64;
 Query Match 95.2%; Score 99; DB 11; Length 99;
 Best Local Similarity 94.7%; Pred. No. 6.5e-08;
 Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NPENSRGKGRGORGKNGR 19
 Db |||||
 Db 73 SPENSRGKGRGORGKNGR 91
 RESULT 6
 O9QXJ9 PRELIMINARY; PRT; 125 AA.
 ID O9QXJ9
 AC Q9QXJ9;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE GDNF splice variant 1 (Fragment).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR; TISSUE=BRAIN;
RA Russell F.D., Koishi K., Jiang Y., McLennan I.S.;
RT "Anterograde axonal transport of glial cell line-derived neurotrophic
RT factor and its receptors in rat hypoglossal nerve.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RD EMBL; AF205713; AAF23767.1; -
RW NON_TER 125
SQ SEQUENCE 125 AA; 13881 MW; 5879E8E5A7190279 CRC64;

Query Match 95.2%; Score 99; DB 11; Length 125;
Best Local Similarity 94.7%; Pred. No. 8.4e-08;
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NPENSRGKGRRGQGRGNRG 19
Db :|||||
Db 99 SPENSRGKGRRGQGRGNRG 117
Db :|||||

RESULT 7
Q8R485
ID Q8R485 PRELIMINARY; PRT; 199 AA.
AC Q8R485;
DT 01-JUN-2002 (TEMBLrel. 21, Created)
DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Neurotrophic factor GDNF (Fragment).
GN GDNF.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21927466; PubMed=11930164;
RA Armstrong K.J., Niles L.P.;
RT "Induction of GDNF mRNA expression by melatonin in rat C6 glioma
RT cells.";
RL Neuroreport 13:473-475(2002).
RD EMBL; AF497634; AM18096.1; -
RW NON_TER 199
SQ SEQUENCE 199 AA; 22299 MW; 0F3D222869386F7D CRC64;

Query Match 95.2%; Score 99; DB 11; Length 199;
Best Local Similarity 94.7%; Pred. No. 1.4e-07;
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NPENSRGKGRRGQGRGNRG 19
Db :|||||
Db 91 SPENSRGKGRRGQGRGNRG 109
Db :|||||

RESULT 8
Q9IAM2
ID Q9IAM2 PRELIMINARY; PRT; 182 AA.
AC Q9IAM2;
DT 01-OCT-2000 (TEMBLrel. 15, Created)
DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TEMBLrel. 20, Last annotation update)
DE Glial cell line-derived neurotrophic factor short form
DE (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20092738; PubMed=10625540;
RA Homma S., Oppenheim R.W., Yaginuma H., Kimura S.;
RT "Expression pattern of GDNF, c-ret, and GFRalpha3 suggests novel roles
RT for GDNF ligands during early organogenesis in the chick embryo.";
RL Dev. Biol. 217:121-137(2000).
RW NON_TER 1
SQ SEQUENCE 199 AA; 22299 MW; 0F3D222869386F7D CRC64;

Query Match 95.2%; Score 99; DB 11; Length 199;
Best Local Similarity 94.7%; Pred. No. 1.4e-07;
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NPENSRGKGRRGQGRGNRG 19
Db :|||||
Db 91 SPENSRGKGRRGQGRGNRG 109
Db :|||||

RESULT 9
Q9IAM3
ID Q9IAM3 PRELIMINARY; PRT; 215 AA.
AC Q9IAM3;
DT 01-OCT-2000 (TEMBLrel. 15, Created)
DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TEMBLrel. 20, Last annotation update)
DE Glial cell line-derived neurotrophic factor long form (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20092738; PubMed=10625540;
RA Homma S., Oppenheim R.W., Yaginuma H., Kimura S.;
RT "Expression pattern of GDNF, c-ret, and GFRalpha3 suggests novel roles
RT for GDNF ligands during early organogenesis in the chick embryo.";
RL Dev. Biol. 217:121-137(2000).
RW NON_TER 1
SQ SEQUENCE 215 AA; 24150 MW; 46A45417EADF8DA0 CRC64;

Query Match 69.2%; Score 72; DB 13; Length 215;
Best Local Similarity 73.7%; Pred. No. 0.0022;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 NPENSRGKGRRGQGRGNRG 19
Db :|||||
Db 71 NVENSSKGRNRGKGRNG 89
Db :|||||

RESULT 10
Q8QGE9
ID Q8QGE9 PRELIMINARY; PRT; 143 AA.
AC Q8QGE9;
DT 01-JUN-2002 (TEMBLrel. 21, Created)
DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Glial cell line-derived neurotrophic factor GDNF (Fragment).
GN GDNF.
OS Nipponia nippon.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```

RT for GDNF ligands during early organogenesis in the chick embryo.";
RL Dev. Biol. 217:121-137(2000).
RC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
RD EMBL; AF176018; AAF26685.1; -
RW HSP; Q07731; IAGO.
RT InterPro; IPR001839; TGFb.
RD Pfam; PF00019; TGF-beta; 1.
RW SMART; SM00204; TGFb; 1.
KW Glycoprotein.
FT NON_TER 182
SQ SEQUENCE 182 AA; 20740 MW; 6A8AC16BD1B4F103 CRC64;

Query Match 69.2%; Score 72; DB 13; Length 182;
Best Local Similarity 73.7%; Pred. No. 0.0018;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 NPENSRGKGRRGQGRGNRG 19
Db :|||||
Db 71 NVENSSKGRNRGKGRNG 89
Db :|||||

RESULT 9
Q9IAM3
ID Q9IAM3 PRELIMINARY; PRT; 215 AA.
AC Q9IAM3;
DT 01-OCT-2000 (TEMBLrel. 15, Created)
DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TEMBLrel. 20, Last annotation update)
DE Glial cell line-derived neurotrophic factor long form (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20092738; PubMed=10625540;
RA Homma S., Oppenheim R.W., Yaginuma H., Kimura S.;
RT "Expression pattern of GDNF, c-ret, and GFRalpha3 suggests novel roles
RT for GDNF ligands during early organogenesis in the chick embryo.";
RL Dev. Biol. 217:121-137(2000).
RW NON_TER 1
SQ SEQUENCE 215 AA; 24150 MW; 46A45417EADF8DA0 CRC64;

Query Match 69.2%; Score 72; DB 13; Length 215;
Best Local Similarity 73.7%; Pred. No. 0.0022;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 NPENSRGKGRRGQGRGNRG 19
Db :|||||
Db 104 NVENSSKGRNRGKGRNG 122
Db :|||||

RESULT 10
Q8QGE9
ID Q8QGE9 PRELIMINARY; PRT; 143 AA.
AC Q8QGE9;
DT 01-JUN-2002 (TEMBLrel. 21, Created)
DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Glial cell line-derived neurotrophic factor GDNF (Fragment).
GN GDNF.
OS Nipponia nippon.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

OC Archosaurola; Aves; Neognathae; Ciconiiformes; Threskiornithidae;

OC Nipponia.

OX NCBI_TaxID=128390;

RN [1]

RP SEQUENCE FROM N.A.

RA Zheng H., Fang S., Xi Y., Fujihara N.;

RT "Cloning and expression of glial cell line-derived neurotrophic factor

(GDNF) of Nipponia nippon."

RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF469665; AAL79041.1; "

FT NON_TER 1 1

FT NON_TER 143 143

SQ SEQUENCE 143 AA; 16507 MW; 26ADB9C00B6231E CRC64;

Query Match

Best Local Similarity 68.3%; Score 71; DB 13; Length 143;

Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 NPENSRGKGRRGGRGKNGR 19

| | | | | | | | | | | | | | |

31 NIENSSKKGRNRQKNGR 49

RESULT 11

Q9LVP1

ID Q9LVP1 PRELIMINARY; PRT; 926 AA.

AC Q9LVP1;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE Arabidopsis thaliana (EC 5.99.1.2).

OS Arabidopsis thaliana (Mouse-ear cross).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.

OX NCBI_TaxID=3702;

OC [1]

RN SEQUENCE FROM N.A.

RP STRAIN=COLUMBIA;

RC MEDLINE=20181125; PubMed=10718197;

RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,

RA Tabata S.;

RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence

features of the regions of 3,076,755 bp covered by sixty P1 and TAC

clones.";

RL DNA Res. 7:31-63(2000).

CC -1- FUNCTION: THE REACTION CATALYZED BY TOPOISOMERASES LEADS TO THE

CONVERSION OF ONE TOPOLOGICAL ISOMER OF DNA TO ANOTHER (BY

SIMILARITY).

-1- CATALYTIC ACTIVITY: ATP-INDEPENDENT BREAKAGE OF SINGLE-STRANDED

DNA, FOLLOWED BY PASSAGE AND REJOINING.

-1- MISCELLANEOUS: WHEN A TOPOISOMERASE TRANSIENTLY BREAKS A DNA

BACKBONE BOND, IT SIMULTANEOUSLY FORMS A PROTEIN-DNA LINK, IN

WHICH A TYROSYL OXYGEN IN THE ENZYME IS JOINED TO A DNA PHOSPHORUS

AT ONE END OF THE ENZYME-SEVERED DNA STRAND (BY SIMILARITY).

-1- SIMILARITY: BELONGS TO PROKARYOTIC TYPE I/III TOPOISOMERASE

FAMILY.

EMBL; AB019227; BAA96895.1; "

InterPro: IPR002936; DnaPrim_toprim.

InterPro: IPR003601; DnaTop_ATP_bind.

InterPro: IPR003602; DnaTop_ATP_bind.

InterPro: IPR000380; Prok_topoisomrase.

InterPro: IPR001878; znf_CCHC.

Pfam: PF01131; Topoisom_bac; 1.

Pfam: PF01751; Toprim; 1.

Pfam: PF01396; zf-C4_Topoisom; 1.

Pfam: PF00098; zf-CCHC; 2.

PRINTS; PR00417; PRTPISMRASEI.

SMART; SM00437; TOP1AC; 1.

SMART; SM00436; TOP1Bc; 1.

SMART; SM00493; TOP1W; 1.

SMART; SM00343; znf_C2HC; 2.

PROSITE; PS00396; TOPOISOMERASE_I_PROK; 1.

KW DNA-binding; Isomerase; Topoisomerase.

SQ SEQUENCE 926 AA; 103407 MW; 806A59EAD08CBB08 CRC64;

Query Match

Best Local Similarity 57.7%; Score 60; DB 10; Length 926;

Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 NSRGKGRRGGRGKNGR 19

| | | | | | | | | | | | | | |

DB 857 NSGSGRGRGSRGRGRG 872

RESULT 12

Q9ATY5

ID Q9ATY5 PRELIMINARY; PRT; 1479 AA.

AC Q9ATY5;

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE UV hypersensitive protein.

GN UVH3

OS Arabidopsis thaliana (Mouse-ear cross).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21332854; PubMed=11439121;

RA Liu Z., Hail J.D., Mount D.W.;

RT "Arabidopsis UVH3 gene is a homolog of the Saccharomyces cerevisiae

RAD2 and human XPG DNA repair genes.";

RL Plant J. 26:329-338(2001).

DR EMBL; AF312711; AAK37472.1; "-.

DR InterPro: IPR000513; Exo_N_I.

DR InterPro: IPR003584; HHH_2.

DR InterPro: IPR001532; XPGC_Rad.

DR Pfam; PF00867; XPG_I; 1.

DR Pfam; PF00752; XPG_N; 1.

DR PRINTS; PR00853; XPGRADSUPER.

DR SMART; SM00279; HHH2; 1.

DR SMART; SM00484; XPGI; 1.

DR SMART; SM00485; XPGN; 1.

DR PROSITE; PS00842; XPG_2; 1.

SQ SEQUENCE 1479 AA; 165666 MW; 72F38D472849EE55 CRC64;

Query Match

Best Local Similarity 55.8%; Score 58; DB 10; Length 1479;

Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 PENSCKGRRGGRGKNGR 19

| : | | | : | | : | | |

DB 1264 PSSRGGRGGRGKNGR 1281

RESULT 13

Q9LRT2

ID Q9LRT2 PRELIMINARY; PRT; 1522 AA.

AC Q9LRT2;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE Similarity to nucleotide excision repair protein.

OS Arabidopsis thaliana (Mouse-ear cross).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=COLUMBIA;

RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;

RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.

```

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=20363099; PubMed=10907853;
RA Nakamura Y.;
RT *Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT TAC and BAC clones.*;
RL DNA Res. 7:217-221(2000).
DR EMBL: AB028616; BAB01125.1; -.
DR InterPro: IPR000513; EXO_N.I.
DR InterPro: IPR003584; HHH2.
DR InterPro: IPR001532; XPGC_Rad.
DR Pfam: PF00867; XPG_I; 1.
DR Pfam: PF00752; XPG_N; 1.
DR PRINTS: PR00853; XPGGRADSUPER.
DR SMART: SM00279; HHH2; 1.
DR SMART: SM00484; XPGI; 1.
DR SMART: SM00485; XPGN; 1.
DR PROSITE: PS00842; XPG_2; 1.
SEQUENCE 1522 AA; 170399 MW; D189F84F4C76CB45 CRC64;

Query Match 55.8%; Score 58; DB 10; Length 1522;
Best Local Similarity 55.6%; Pred. No. 2.5;
Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 PENSRGKRRGQKNGRG 19
DB 1304 PSSSRGGRGAQKRGGRG 1321

RESULT 14
Q8RZE5 PRELIMINARY; PRT; 513 AA.
AC Q8RZE5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative AAA-type ATPase-like protein.
GN P0510C12.7.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoidae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
PP SEQUENCE FROM N.A.
STRAIN=CV. NIPPONBARE;
Sasaki T., Matsumoto T., Yamamoto K.;
RT *Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0510C12.*;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF003725; BAB90463.1; -.
SQ SEQUENCE 513 AA; 55589 MW; 05CDDBE17327314A CRC64;

Query Match 54.8%; Score 57; DB 10; Length 513;
Best Local Similarity 64.7%; Pred. No. 1.1;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 ENSRGKRRGQKNGRG 19
DB 495 EISHGQGRGRRGVRGGRG 511

RESULT 15
Q8S6K4 PRELIMINARY; PRT; 545 AA.
AC Q8S6K4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 62.7 kDa protein.
GN OSJNBA0019N10.22.

```

```

OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoidae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
PP SEQUENCE FROM N.A.
STRAIN=NIPPONBARE;
McCombie W.R., de la Bastide M., Spiegel L., Preston R., Kirchoff K.,
RA Kuit K., Nascimento L., Zutavern T., Ballja V., Bell M., Baker J.,
RA Santos L., Miller B., Katzenberger F., Muller S., King L., Yang C.,
RA O'Shaughnessy A., Palmer L., Dedhia N.;
RT *Genomic sequence for Oryza sativa, Nipponbare strain, clone
RT OSJNBA0019N10, from chromosome 10, complete sequence.*;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC092748; AA01084.1; -.
DR Hypothetical protein.
SQ SEQUENCE 545 AA; 62654 MW; 799E292BE497DC90 CRC64;

Query Match 53.4%; Score 55.5; DB 10; Length 545;
Best Local Similarity 50.0%; Pred. No. 2;
Matches 14; Conservative 2; Mismatches 3; Indels 9; Gaps 2;

QY 1 NPEN-----SRGKRRGQ-RGKNGRG 19
DB 482 NPANFEKVRKAMSRGRRGRPRGSGRG 509

Search completed: December 4, 2002, 08:18:53
Job time : 32 secs

```

THIS PAGE BLANK (USPTO)


```
RESULT 2
US-08-535-681-19
; Sequence 19, Application US/08535681
; Patent No. 6184200
; GENERAL INFORMATION:
; APPLICANT: Hu, Sylvia
; TITLE OF INVENTION: Truncated Glial Cell Line-Derived Neurotrophic
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: United States of America
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/535,681
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Curry, Daniel R.
; REGISTRATION NUMBER: 32,727
; REFERENCE/DOCKET NUMBER: A-357
; TELEPHONE: 805-447-8102
; TELEFAX: 805-499-8011
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-535-681-19
Query Match 100.0%; Score 104; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 4e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NPNSRGKGRGGRGKNG 19
DB 2 NPNSRGKGRGGRGKNG 20
|||||
RESULT 3
US-08-535-681-20
; Sequence 20, Application US/08535681
; Patent No. 6184200
; GENERAL INFORMATION:
; APPLICANT: Hu, Sylvia
; TITLE OF INVENTION: Truncated Glial Cell Line-Derived Neurotrophic
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: United States of America
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/535,681
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Curry, Daniel R.
; REGISTRATION NUMBER: 32,727
; REFERENCE/DOCKET NUMBER: A-357
; TELEPHONE: 805-447-8102
; TELEFAX: 805-499-8011
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-535-681-19
Query Match 100.0%; Score 104; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 4e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NPNSRGKGRGGRGKNG 19
DB 2 NPNSRGKGRGGRGKNG 20
|||||
RESULT 3
US-08-535-681-20
; Sequence 20, Application US/08535681
; Patent No. 6184200
; GENERAL INFORMATION:
; APPLICANT: Hu, Sylvia
; TITLE OF INVENTION: Truncated Glial Cell Line-Derived Neurotrophic
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: United States of America
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/535,681
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Curry, Daniel R.
; REGISTRATION NUMBER: 32,727
; REFERENCE/DOCKET NUMBER: A-357
; TELEPHONE: 805-447-8102
; TELEFAX: 805-499-8011
; TELEX:
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-535-681-21
Query Match 100.0%; Score 104; DB 4; Length 22;
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/535,681
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Curry, Daniel R.
; REGISTRATION NUMBER: 32,727
; REFERENCE/DOCKET NUMBER: A-357
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 805-447-8102
; TELEFAX: 805-499-8011
; TELEX:
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-535-681-20
Query Match 100.0%; Score 104; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.2e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NPNSRGKGRGGRGKNG 19
DB 3 NPNSRGKGRGGRGKNG 21
|||||
RESULT 4
US-08-535-681-21
; Sequence 21, Application US/08535681
; Patent No. 6184200
; GENERAL INFORMATION:
; APPLICANT: Hu, Sylvia
; TITLE OF INVENTION: Truncated Glial Cell Line-Derived Neurotrophic
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: United States of America
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/535,681
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Curry, Daniel R.
; REGISTRATION NUMBER: 32,727
; REFERENCE/DOCKET NUMBER: A-357
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 805-447-8102
; TELEFAX: 805-499-8011
; TELEX:
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-535-681-21
Query Match 100.0%; Score 104; DB 4; Length 22;
```

```
Best Local Similarity 100.0%; Pred. No. 4.4e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NPENSRGKRRGGRGKNGR 19
    |||||
Db 4 NPENSRGKRRGGRGKNGR 22

RESULT 5
US-08-535-681-22
; Sequence 22, Application US/08535681
; Patent No. 6184200
; GENERAL INFORMATION:
; APPLICANT: Hu, Sylvia
; TITLE OF INVENTION: Truncated Glial Cell Line-Derived Neurotrophic
; FACTOR
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: United States of America
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/535,681
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Curry, Daniel R.
; REGISTRATION NUMBER: 32,727
; REFERENCE/DOCKET NUMBER: A-357
; TELEPHONE: 805-447-8102
; TELEFAX: 805-499-8011
; TELEX:
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-535-681-22

Query Match 100.0%; Score 104; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 4.6e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NPENSRGKRRGGRGKNGR 19
    |||||
Db 5 NPENSRGKRRGGRGKNGR 23

RESULT 6
US-08-535-681-23
; Sequence 23, Application US/08535681
; Patent No. 6184200
; GENERAL INFORMATION:
; APPLICANT: Hu, Sylvia
; TITLE OF INVENTION: Truncated Glial Cell Line-Derived Neurotrophic
; FACTOR
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
```

```
COUNTRY: United States of America
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/535,681
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Curry, Daniel R.
REGISTRATION NUMBER: 32,727
REFERENCE/DOCKET NUMBER: A-357
TELEPHONE: 805-447-8102
TELEFAX: 805-499-8011
TELEX:
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-535-681-23

Query Match 100.0%; Score 104; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 4.8e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NPENSRGKRRGGRGKNGR 19
    |||||
Db 6 NPENSRGKRRGGRGKNGR 24

RESULT 7
US-08-535-681-24
; Sequence 24, Application US/08535681
; Patent No. 6184200
; GENERAL INFORMATION:
; APPLICANT: Hu, Sylvia
; TITLE OF INVENTION: Truncated Glial Cell Line-Derived Neurotrophic
; FACTOR
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: United States of America
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/535,681
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Curry, Daniel R.
; REGISTRATION NUMBER: 32,727
; REFERENCE/DOCKET NUMBER: A-357
; TELEPHONE: 805-447-8102
; TELEFAX: 805-499-8011
; TELEX:
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
```

APPLICANT: Hu, Sylvia
TITLE OF INVENTION: Truncated Glial Cell Line-Derived Neurotrophic

TELECOMMUNICATION INFORMATION:
TELEPHONE: 805-447-8102
TELEFAX: 805-499-8011
TELEX:
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-535-681-27

Query Match 100.0%; Score 104; DB 4; Length 28;
Best Local Similarity 100.0%; Pred. No. 5.6e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NPENSRGKGRGGRGKNGR 19
|||||
10 NPENSRGKGRGGRGKNGR 28

RESULT 11
US-08-535-681-28
Sequence 28, Application US/08535681
Patent No. 6184200
GENERAL INFORMATION:
APPLICANT: Hu, Sylvia
TITLE OF INVENTION: Truncated Glial Cell Line-Derived Neurotrophic
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: AMGEN INC.
STREET: 1840 DeHavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: United States of America
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/535,681
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Curry, Daniel R.
REGISTRATION NUMBER: 32,727
REFERENCE/DOCKET NUMBER: A-357
TELECOMMUNICATION INFORMATION:
TELEPHONE: 805-447-8102
TELEFAX: 805-499-8011
TELEX:
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-535-681-28

Query Match 100.0%; Score 104; DB 4; Length 29;
Best Local Similarity 100.0%; Pred. No. 5.8e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NPENSRGKGRGGRGKNGR 19
|||||
11 NPENSRGKGRGGRGKNGR 29

RESULT 12
US-08-535-681-29
Sequence 29, Application US/08535681
Patent No. 6184200
GENERAL INFORMATION:
APPLICANT: Hu, Sylvia
TITLE OF INVENTION: Truncated Glial Cell Line-Derived Neurotrophic
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: AMGEN INC.
STREET: 1840 DeHavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: United States of America
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/535,681
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Curry, Daniel R.
REGISTRATION NUMBER: 32,727
REFERENCE/DOCKET NUMBER: A-357
TELECOMMUNICATION INFORMATION:
TELEPHONE: 805-447-8102
TELEFAX: 805-499-8011
TELEX:
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-535-681-29

Query Match 100.0%; Score 104; DB 4; Length 30;
Best Local Similarity 100.0%; Pred. No. 6e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NPENSRGKGRGGRGKNGR 19
|||||
Db 12 NPENSRGKGRGGRGKNGR 30

RESULT 13
US-08-535-681-30
Sequence 30, Application US/08535681
Patent No. 6184200
GENERAL INFORMATION:
APPLICANT: Hu, Sylvia
TITLE OF INVENTION: Truncated Glial Cell Line-Derived Neurotrophic
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: AMGEN INC.
STREET: 1840 DeHavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: United States of America
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

Query Match 100.0%; Score 104; DB 4; Length 32;
Best Local Similarity 100.0%; Pred. No. 6.4e-09;

```
Query Match          100.0%; Score 104; DB 4; Length 32;
Best Local Similarity 100.0%; Pred. No. 6.4e-09;
```

```
QY      1 NPENSRCKGRRGORCKNRG  19  
        |||||  
Db      15 NPENSRCKGRRGORCKNRG  33
```

Search completed: December 4, 2002, 08:19:53
Job time : 16 secs

Search completed: December 4, 2002, 08:19:53
Job time : 16 secs